08

E7.4-10687. III
CR-139220

21 June 1974

FINAL REPORT
ON
AUTOMATIC PHOTOINTERPRETATION FOR LAND USE
MANAGEMENT IN MINNESOTA

(E74-10687) AUTOMATIC PHOTOINTERPRETATION FOR LAND USE MANAGEMENT IN MINNESOTA Final Report (Honeywell, Inc.) 65 p HC $6.25 CSCL 08B
N74-32760
Unclas
G3/13 00687

ERTS-1
Proposal Number MMC 257
Principal Investigator Number PR 202

Contract Number NAS5-21742

Prepared by: L. Kirvida
M. Cheung
D. Pile
R. Zirkle

Approved by: G. D. Swanlund by SS
G. D. Swanlund

Honeywell, Inc.
Systems and Research Division
2600 Ridgway Parkway
Minneapolis, Minnesota 55413

Figure		Page

Final Report

Automatic Photointerpretation for Land Use Management in Minnesota

Abstract

Automatic Photointerpretation techniques were utilized to evaluate the feasibility of data for land use management. It is shown that ERTS MSS data can produce thematic maps of adequate resolution and accuracy to update Land Use Maps. In particular, five typical Land Use areas were mapped with classification accuracies ranging from 77% to over 90%.

I. Introduction

Our major objective in this study has been to apply automatic photointerpretation techniques to land use thematic mapping from ERTS MSS data. The land use classes selected were guided by those used in the 1969 Minnesota State Land Use Map, i.e. forested, cultivated, open, water, marsh, urban residential, urban non-residential, extraction and transportation. The 1969 map was prepared from aerial photographs and manual interpretation.

The study approach was a set of experiments using ERTS MSS data. Five geographical areas in Minnesota were chosen for analysis. A multiclass (K-Class) automatic classification procedure was utilized to prepare thematic maps. The fifth experiment consisting of a lake count was performed by setting an upper threshold on band 7 for detecting water bodies.

In one of the experiments, the automatic classifier was trained on a training area and tested on another area approximately 25 miles from the training site. This experiment was performed to determine the generality of training weights. There was no noticeable degradation in performance over this distance.

The addition of texture features was evaluated on the metropolitan area. Texture improves classification performance

but its measurement requires larger areas of the same class. Thus the ground resolution is degraded.

The quality of the input data was generally good. In several of the images, one of the six channels either put out an attenuated signal or experienced random drop-outs. To accommodate this error source, these channels were delt with separately, i.e. separate classification weights were determined for any channel whose output was noticeably different from the norm.

II. Experimental Design

The data used for constructing the feature vector was obtained from 7-track 800 BPI computer compatible tapes. The bulk data on black and white 9.5 inch positive transparencies was used to locate the area of interest. The ERTS 9½ inch bulk photos, the RB 57 aerial photos, county maps and Mark Hurd aerial survey photos were used to select the areas to be analyzed. These data blocks were then extracted from the four computer compatible tapes (CCT). For most of the small data blocks, only one digital tape was required. When two tapes were needed to cover the desired area, alignment of data between the two tapes was achieved by aligning data having the same record number.

The extracted data block from the CCT was reformatted for writing out on an Optronics film writer or on a line printer which printed out a grey map. For writing out the image, the dynamic range of the film writer was matched to the density values of the image as determined from a histogram plot. The Optronics films were written out in 64 equal density level increments. The line printer grey maps were composed of over-writing four alpha numeric characters to achieve 16 density levels. These levels were also assigned on the basis of the intensity histograms. An attempt was made to achieve different density symbols for the various classes.

The selection of classes was guided by the Minnesota Land Management Information System (MLMIS) classes. Of the nine classes used in this system,seven were included in this study. Cultivated regions were not considered because of the unavailability of current ground truth. In addition, the transportation class was not broken out as a separate class because of the lack of training samples.

Data Analysis Procedure

After receipt of the ERTS data on photographs and on digital tapes, the photographic images were checked for gross anomolies in the detector operation. When banding was apparent, histo-

grams were computed for each of the six detector scan lines to detect differences in the mean and variance in the output of each detector. When large differences occurred in the output of the detector channels, they were processed separately.

The ERTS data was registered with available ground truth (obtained from aircraft photography and sample ground inspection) by recognition of landmarks. For example, lakes on MSS band seven provided good landmarks for registering the two images. An enlarged view of the ERTS image landmarks were obtained by converting the ERTS digital tape to film with an Optronics Digital tape-to-film writer. The film record output is an image containing grid lines corresponding to record and word numbers on the CCT. In addition, the landmarks are displayed in context on the image, thus simplifying photointerpretation and registration. This procedure allows registering the tape to a known landmark to within a single resolution element. Registration of the ERTS data with ground-truth maps is essential for both training and evaluating the automatic interpretation system and for producing the stratification overlay. Registration is also essential for adding successive coverage data.

Once ground truth and ERTS data are registered, class boundaries can be encoded in terms of record and word numbers on the digital tape. From within these class boundaries, the training samples were extracted. The generation of the thematic maps was achieved by using a trained classifier. Training is performed by a "teacher", i.e. a training set of data is isolated for each of the classes of interest. The training set is used to generate decision weights as follows.

Samples of each class from the training set are grouped together in a least-mean-square sense about points in decision space which are equidistant from each other and from the origin. Linear boundaries are determined to minimize the error assignments for this training set. These boundaries or decision weights are then used to categorize the input data during the formation of the thematic map. While generating these decision boundaries, a confusion matrix is printed. This provides performance data for the operator to use in evaluating the expected accuracy of classification.

By using the spectral distribution of energy over the classes of interest, it is evident that some of the classes are separable. For example, the histograms in Figure 1 indicate

that water and conifers are easy to identify on this October image taken over Cloquet, Minnesota. Because of the leaf-off condition, hardwoods are not as easy to separate from open areas and cities. Adding a January coverage greatly improves separability.

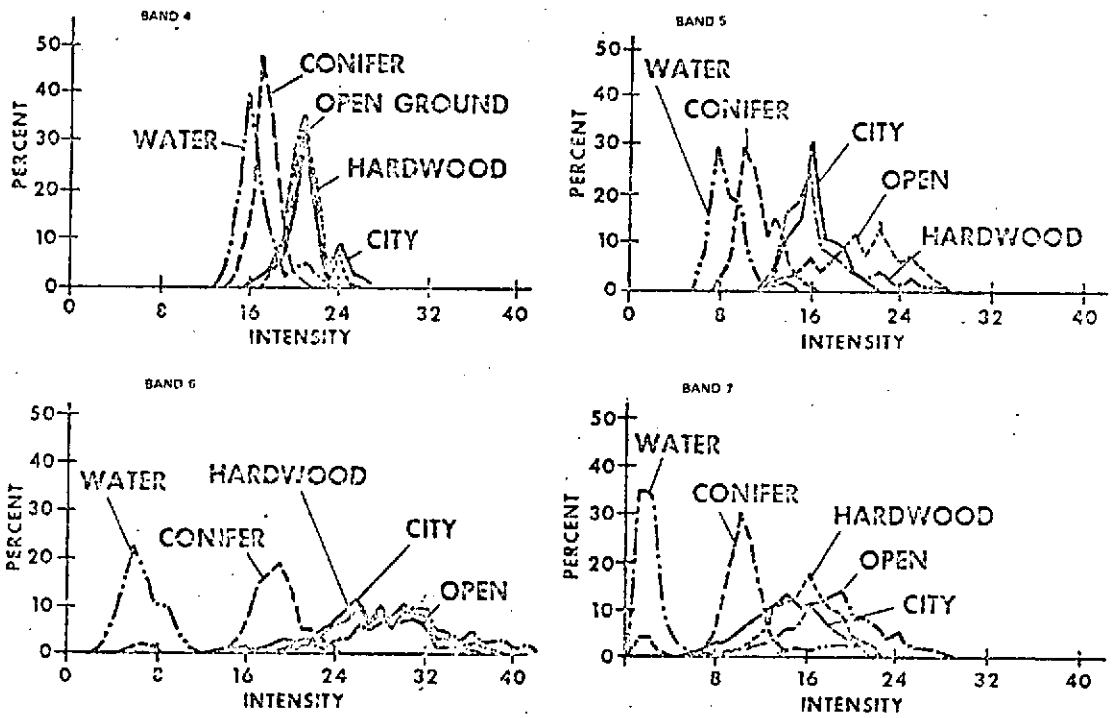


Figure 1. Histograms of Intensity Levels
ERTS 1075-16312

A block diagram depicting the chain of events in the automatic photointerpretation procedure is shown in Figure 2. The following discussion is centered around Figure 2. The data block from the 100 by 100 mile ERTS image will be selected by longitude and latitude specifications. Finer resolution can be obtained by using aerial photos and landmarks for depicting the desired area. Because of the earth's rotation, the ERTS photos are not aligned to the north despite the polar orbit. This requires a slightly larger rectangle of data to include all of the area described by corner points expressed in terms of longitude and latitude.

The ERTS bulk tapes are then reformated to be compatible with a film writer. Data from each of the bands is written out using 6 bits per byte. The data is also reformated so that intensity levels from the four bands for each pixel are adjacent to one another for the automatic classifier feature vector. One channel of the selected data block is then written out on a film writer and registered with ground truth. If dual coverages are used, registration is made with a second coverage. Registration between dual coverages (October and January) over Northern Minnesota required translation in longitude and latitude and a rotation of 0.03 degrees of one coverage with respect to the other.

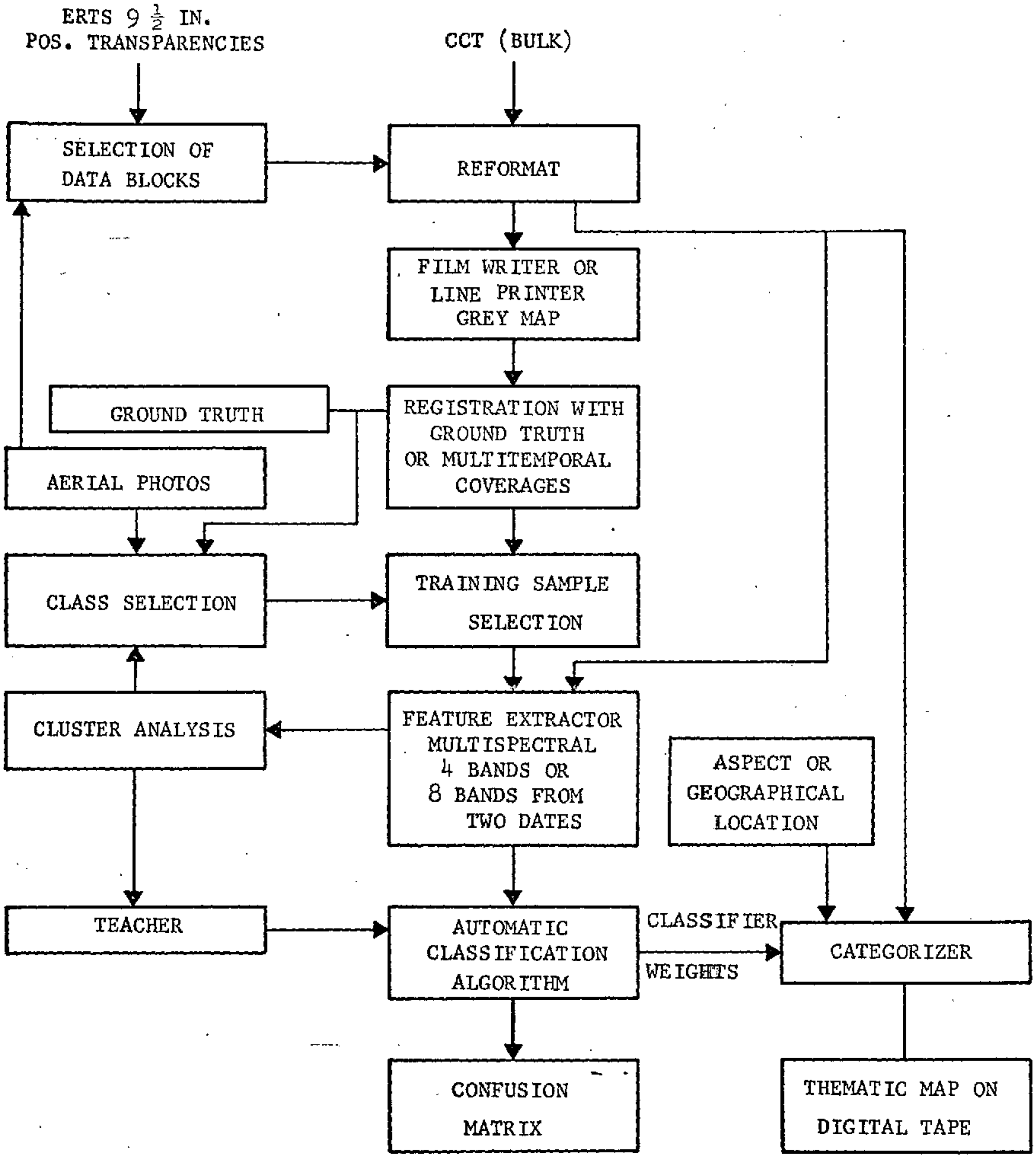


Figure 2. Automatic Photointerpretation Procedure

The classes selected should be guided by separability, that is, by some insight into the natural clustering of the data. Once the classes have been chosen, training samples are selected from representative areas to determine classifier weights.

The K-Class training procedure generates a confusion matrix, which permits judging the performance. The confusion matrix lists the actual ground-truth-versus-classifier assignments for all samples, thus one can determine where the most errors are made. A cost procedure can also be used for penalizing certain misclassifications, thus reducing this type of error.

After obtaining an acceptable classification performance on the training set, these classifier weights are used to categorize the selected blocks of data. Ancillary information, such as aspect or geographical location, could be fed into the categorizer at the same time to improve classification accuracy. The output of the categorizer is a digital magnetic tape with each pixel coded into one of the prescribed classes.

The color coded thematic map is generated from the thematic map on the digital data tape by writing out the coded pixels into three black and white transparencies. The transparencies can be formed such that any class can be assigned to any desired color. The color spectrum can be divided into any required number of colors, however as the required number

of colors increases, the change in hue between classes decreases. The three primary color black and white transparencies are then color combined into the final color-coded thematic map.

III. Experimental Results

Ramsey County Lake Study

This study was run using Ramsey County as the test area due to the availability of an ERTS coverage and easily accessible ground truth data. The ERTS photo 1075-16321, an October 6, 1972 coverage was used. Lake acreages were determined by setting a threshold on band 7 and counting areas below this threshold level. Open water generally provides a signal of 5 or less, marsh areas can be detected in the range of 6-9.

An inventory of water resources is valuable in determining water resources, recreational areas, and storage capabilities. Presently, this is done from aerial photos which are costly, time consuming, and, therefore, not taken often. Potentially, ERTS photos offer an 18-day inventory updating system. This information could also be used to improve the water information on county highway maps since it is not unusual for existence, size, location, and/or shape of a lake on these maps to be erroneous.

More rapid changes can be observed by repetitive coverage such as that provided by ERTS. These involve changes in shoreline and lake size due to natural processes and development, or the side effects of development. The resulting changes can effect waterfowl, fish and also legal matters; eg., the bringing of unused shallow lake lands into productive use involves a taxation change on that land.

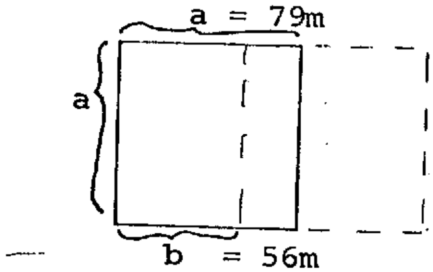
The monitoring of changes in lake levels and lake occurrences is also of great value in developing an index of annual waterfowl production. Canadian and U. S. biologists suggested that a reliable production index could be derived independent of size of the breeding population from estimates of the number of ponds remaining in mid-July.* Automatic data processing of changes in wetness during the May-July period would provide this.

For this study, a print out of the band 7 was used. Since each data point is approximately equal to 1.5 acres on the ground, one had only to count up the groupings of numbers and multiply the obtained total by a correction factor. To account for element overlap along the scan lines, the expression 1.104N + .453L was used. In the expression, N is the number of elements and L is the number of scan lines involved in the composition of any irregular shaped object.

---

*"Preliminary Evaluation of ERTS-1 for Determining Numbers and Distribution of Prairie Ponds and Lakes"; Edgar A.Work, Jr., Environmental Research Institute of Michigan; David S. Gilmer and A. T. Klett, Bureau of Sport Fisheries and Wildlife Northern Prairie Wildlife Research Center.

The derivation of this relationship is shown below:



Area of one strip = a [a + (n-1) b]

$$\text{Area of L strips} = \sum_{\ell=1}^{L} a[(a-b) + n_\ell b]$$

$$= L\,a(a-b) + ab \sum_{\ell=1}^{L} n_\ell$$

$$= L\,a(a-b) + ab\,N = (1817L + 4424N) \text{ meter}^2$$

$$A = 0.453L + 1.104N \quad \text{acres}$$

After comparing the shape of several lakes on the computer print out with their respective shape on topographic maps, the interval of 0 through 5 on band 7 was selected as being indicative of water. The acreage of each lake in Ramsey County was then counted and compared with available data on basin and lake acreage.

It should be pointed out that the interval selection is somewhat arbitrary. Any error in the choice of an interval should produce consistant errors in acreage, i.e., consistantly too high or consistantly too low.

The acreages determined from the ERTS data were compared with available data. Any lakes in Ramsey County not included in this study did not fall within the area pulled from the ERTS photo.

A glance at the counted acreage as compared to basin acreage and lake acreage reveals that, in most cases, the counted acreage provides a good measure of the actual acreage. The comments and lake types included are an attempt to account for the discrepancies that arise.

In comparing the figures, one should bear in mind that basin acreage indicates the area of the basin, i.e., the depression, not the actual area that contains water. Likewise, the lake acreage includes most marsh areas around the lake, not just areas of open water. The chosen interval of 0 through 5, on the other hand, seems to indicate only open water, or water with a very slight amount of emergent vegetation or rushes. This can be seen by looking at some of the lakes indicated as eliminated by air photo in the comment column.

The phase "eliminated by air photo" indicates that a particular lake was considered to contain no open water, i.e., to be drained or swamp, from analysis of 1968 low altitude air photos of Ramsey County. As can be noted from the table, several of the eliminated lakes are shown as containing small amounts

of water. These are likely lakes that contain water in only part of the year and are swamp or dry for the rest. Some of the lakes, for example 62-20, in the upper right quadrant of Figure 3, can be made to assume their exact basin shape by considering the numbers 0 through 9. This seems to imply that the lake is surrounded by detectable swamp (this lake is shown as surrounded by swamp or marsh on the USGS topographic sheet).

Other lakes, for example 62-23, indicated as swamp from field checking, are not visable at all until the interval 6 through 9 is noted. In addition, an area just north of Pigs Eye, 62-4, shows up in the interval 6 through 9 and is designated swamp on the topographic sheet. It seems, then, that this interval, 6 through 9, could be used to identify swamps.

This technique is not useful for determining depth in lakes. While the data shows that many of the counted and planimetric calculations are close, this similarity is coincidental. A comparison of the areal distribution of lake acreage having a depth of 10 feet or greater, as determined from the Division of Water, Soil, and Minerals map and this same acreage determined from the computer print out, shows no correlation. This lack of correlation is likely due to such factors as sun angle and lake water pollution.

Lake No.	Name	Counted Acreage[1]	Basin Acreage[2]	Lake Acreage[3]	Acres Depth ≥ 10 feet[4]	Counted Acres Depth ≥ 10 feet[5]	Type[6]	Comments[7]
62-1	Silver	60.68	68	73			V	seems to match exactly the 5' contour
62-4	Pigs Eye	547.91	511				IV	much swamp to east & south
62-5	Caseys	8.02	14				VA	basin slightly larger, mainly open water
62-6	Kohlman	72.07	84				VA	basin (depression) larger
62-7	Gervais	205.07	234	206	135	138.75	V	basin larger than lake sunf
62-8	Savage	17.32	46				IV	I35E now cuts through lake; basin much larger than water area reeds & emergent veg.
62-9	Round	10.24	14				VA	basin larger than actual lake
62-10	Keller	70.93	72				VA	basin larger
62-11	Wakefield	15.58	23				VA	1/5 shoreline is marsh; basin slightly larger
62-12	Round	18.01	23				V	marsh in area to north that appears to be basin
62-13	Phalen	191.09	193	192	124	127.65	V	basin slightly larger
62-14	no name	6.91	18				VA	basin larger than water area; marshes to NW & SE

Lake No.	Name	Counted Acreage	Basin Acreage	Lake Acreage	Acres Depth ≥ 10 feet	Ac[redacted]ted Acres Depth ≥ 10 feet	Type	Comments
62-15	Sandy	not visable	22				VA	sludge basin, St. Paul water supply
62-16	Beaver	68.91	65				VA	marsh to NE; small bit to south
62-17	no name	17.35	24				IV	marshy & swampy
62-19	Black	9.78	120				IV	surrounded by marsh; eliminated by air photo
62-20	no name	3.78	22				IV	eliminated by air photo
62-21	Tammarack	1.56	69				III	eliminated by air photo
62-22	no name	1.56	10				VA	marsh; inaccessible
62-23	no name	not visable	14				VA	swamp
62-24	Birch	107.56	127				VA	96-county road G-now goes through south end
62-25	Ox	7.36	13				VA	pond, swamp, some open water; 96-county rd. G goes through
62-26	no name	5.80	12				IV	southwestern & northwestern shores-marsh, narrow strip connecting top & bottom
62-27	Gillfillan	89.80	87				VA	
62-28	Sucker	56.70	59	60	12	25.53	V	St. Paul water supply
62-29	Basswood	not vis.	110				III	eliminated by air photo

Lake No.	Name	Counted Acreage	Basin Acreage	Lake Acreage	Acres Depth ≥ 10 feet	C[redacted]ted Acres Depth ≥ 10 feet	Type	Comments
62-30	Lambert	not vis.	542				III	eliminated by air photo
62-31	Grass	not vis.	84				III	eliminated by air photo
62-32	Rice	1.563	121				III	eliminated by air photo
62-33	no name	not vis.	87				IV	eliminated by air photo
62-34	Goose	122.64	152				VA	US 61 goes through
62-35	no name	2.67	10				IV	eliminated by air photo
62-36	Priebe	3.13	17				VA	narrow; surrounded by houses now; used as neighborhood pond
62-37	Gem	16.24	20				VA	marsh at SE end, basin slightly larger than lake
62-38	Vadnais	568.52	477				V	marsh in SW & NE has wate on printout, also water separate from lake in marsh; road there too narrow to show; St.Paul Water
62-39	Twin	25.12	37	35	25	12.21	V	some marsh to SE & NE
62-40	Willow	30.47	75				IV	surrounded by marsh,basin (depression)much larger than water surface
62-41	no name	8.02	15				VA	surrounded by marsh; very shallow, some construc-tion one side & homes; basin large

Lake No.	Name	Counted Acreage	Basin Acreage	Lake Acreage	Acres Depth ≥ 10 feet	Co[redacted]ed Acres Depth ≥ 10 feet	Type	Comments
62-42	no name	6.91	12				VA	basin slightly larger
62-45	Long	4.42	111				III	eliminated by air photo
62-46	Pleasant	553.88	585	627	369	376.29	V	marsh at northwestern edg
62-48	Bennett	19.57	41				III	surrounded by marsh; eliminated by air photo
62-49	Lanton	16.28	35				IV	surrounded by rushes, some open water, emergent veg.
62-50	Wilson	1.56	19				IV	eliminated by air photo
62-51	Robinson	3.78	28				II	eliminated by air photo
62-52	Poplar	1.56	19				IV	eliminated by air photo
62-53	no name	9.13	15				VA	exit ramp from route 36 to route 51 cuts into; some marsh, golf course on cast, housing to north
62-54	McCorron	62.20	71	70	50	44.40	V	a little marsh in NW, basin slightly larger
62-55	Como	63.36	69	72	18	7.77	VA	basin slightly larger than water surface
62-56	Owasso	333.28	360	355	101	241.98	V	marsh at southern and western end; development
62-57	Josephine	112.20	110	115	51	69.93	V	
62-58	Little Johanna	16.24	18				V	surrounded by marsh

Lake No.	Name	Counted Acreage	Basin Acreage	Lake Acreage	Acres Depth ≥ 10 feet	Counted Acres Depth ≥ 10 feet	Type	Comments
62-59	Marsden	1.56	291				IV	eliminated by air photo
62-61	Turtle	418.46	444	447	216	330.78	V	
62-62	Charley	29.56	31				VA	
62-63	no name	3.13	32				II	eliminated by air photo
62-64	Martha	10.24	34				IV	eliminated by air photo
62-65	Sunfish	7.36	12				VA	eliminated (enclosed by Twin Cities Arsenal)
62-70	Round	100.90	122				V	
62-71	Valentine	50.48	58	60	5	16.65	VA	constriction to approx. 300 ft. at northern edg
62-72	Karth	9.58	15				VA	swamp, appears to have been partially filled i
62-73	Snail	138.42	195	185	25	88.80	V	small separate section in NW visable on print out (is very shallow w/much veg.)
62-74	Grass	1.56	146				IV	eliminated by air photo
62-75	Island	44.29	63				VA	4 lane highway (694) put through upper secti some of lake taken
62-76	Jones	8.02	28				IV	emergent veg., only small area of open water, marsh all around

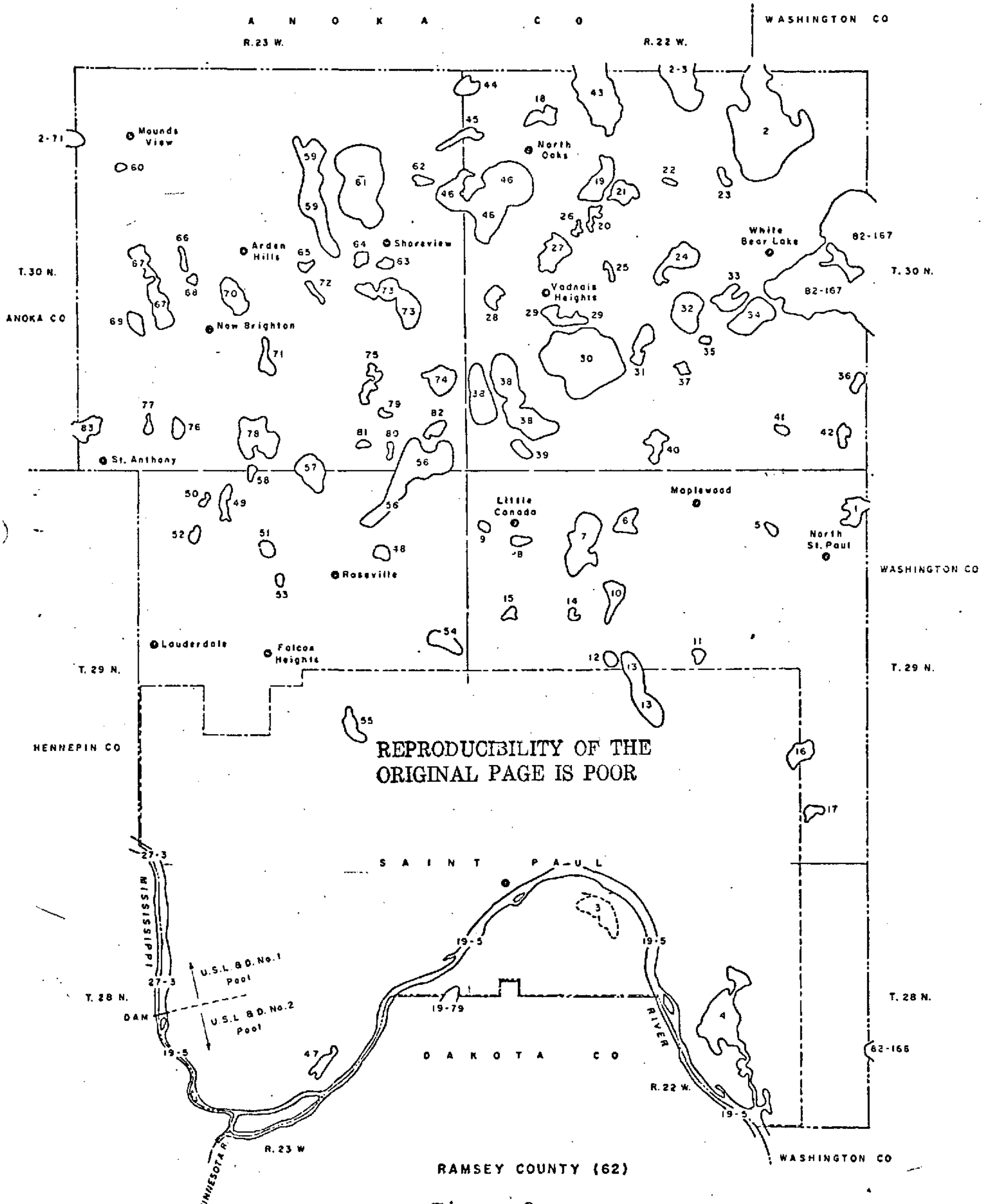
Lake No.	Name	Counted Acreage	Basin Acreage	Lake Acreage	Acres Depth ≥ 10 feet	Cou[redacted]d Acres Depth ≥ 10 feet	Type	Comments
62-78	Johanna	189.49	211	225	162	143.19	V	
62-79	no name	6.46	11				VA	pond
62-80	Emily	5.80	12				VA	highway along eastern length houses on western
62-81	Mud	9.13	16				VA	surrounded on western half by houses & rushes on eastern
62-82	Wabasso	36.67	47	46	20	17.76	V	surrounded by marsh (or rushes) about 50 yds. to open water
19-79	Pickerel	94.94	52				IV	steep side to SE (causes shadow marsh to S & NW toward river; RR through the top
82-167	White Bear	6350.363 2383.57	10965 2410	2479	1924	2065.71	V	

For location of lakes consult map on page 9.

1. number of points on computer printout in the range 0 through 5 multiplied by 1.104N + .453L
2. acreage of lake basin as indicated by Bull. 25, Div. of Water, Soils & Minerals, Dept. of Natural Resources, from 1949 data
3. acreage by planimetry from Div. of Water, Soils, & Minerals maps ranging in age from lake to lake
4. acreage by planimetry from Div. of Water, Soils, & Minerals maps ranging in age from lake to lake
5. number of points on computer print out in the range 0 through 1 multiplied by 1.104N + .453L
6. Lake type established by Metropolitan Inventory of Lakes, Dept. of Conservation, Fish & Game Commission

   II - grass on sedge meadows   IV - deep marsh   VA - marginal fish - game lake
   III - shallow marsh   V - fish lake
7. comments refer to field checking and air photo interpretation (1968 - Mark Hurd photos) done for Ramsey County Lake Surface Zoning Study and reference to topographic maps (USGS - various years)

MINNESOTA CONSERVATION DEPARTMENT



RAMSEY COUNTY (62)

Figure 3

One problem encountered in this study was that things other than water fell in the interval 0 through 5. Along with water were cloud shadows and heavily industrialized areas. It seems, however, that consideration of band 5, which enhances cultural features, will enable the elimination of those areas in the interval 0 through 5 which are not water. On band 5, the heavily industrialized areas seem to show up as much lower numbers than the lakes. Cloud shadows can be distinguished by associating them with clouds on band 5. By looking in the appropriate direction, given the time of day of the photo and the time of year, one encounters an area, somewhat smaller than the cloud shadow, of very high numbers, higher than any thus far encountered in this study. Thus, by scanning the photo, cloud shadows can be distinguished from water. Alternatively, photos from different dates could be compared to detect the presence of clouds.

Sandy Lake, 62-15, was not visable. It is likely that 62-15 was nearly dry at the time of the photo; it is a water supply lake, and therefore highly fluctuating.

It seems that lakes down to a two acre size can be located by this technique. This is a much better performance than that obtained by a visual mapping project conducted by the Department of Geography of the University of Minnesota. They had approxi-

mately a ten acre threshold viewing size for small water bodies. They also noted difficulties in the discrimination of lakes and cloud shadows and of lakes and freshly plowed fields.* The latter difficulty was not encountered in this study due to the lack of agriculture in Ramsey County.

In conclusion, it seems that the technique of determining and monitoring lake acreage would provide adequate information for the various purposes discussed earlier.

Numerous water areas not designated as lakes, perhaps due to their periodic nature, were noted on the computer print out. These corresponded to areas of water or swamp on the USGS topographic maps. Since the appearance of lakes is a reflection of the ground water level, water fluctuation in a swamp area of known elevation could be used to calculate the water table fluctuation. This fluctuation may be due to natural causes or human intervention. This information, when combined with other information dealing with inflow and outflow from a lake by surface systems, would enable the determination of mass budgets for lakes.

---

*"Application of ERTS-1 Imagery of State-Wide Land Information System in Minnesota ", ERTS-1 Application to Classification and Mapping of Water Resources in Minnesota; Dwight A. Brown, Ralph Sanders, Jack Flynn, and John Harrington, Department of Geography, University of Minnesota.

This acreage calculation method could also be valuable in the typing of lakes. A comparison of counted acreage with basin acreage and counted acreage from 6 through 9 could lead to a useful classification of lakes. Jack Flynn, working on the Ramsey County Lake Surface Zoning Study for the Minnesota Land Management Information Systems Study, felt that this computerized evaluation technique would be of great value to the classification of lakes. Some lakes, not eliminated by air photo, were discovered to be marsh on field checking. Others were discovered to be inaccessible due to being completely surrounded by marsh. A computerized evaluation using ERTS-data rather than outdated aerial photos could have led to more efficient use of costly, time-consuming field work.

Trout Lake Area Study

This area was chosen for analysis because of the availability of an ERTS cloud-free coverage and the availability of ground truth. The ERTS frame used for the data source was 1075-16312, an October 6th overpass. The ground truth was obtained from aerial photos, county maps, quadrangle maps and a land use map produced by the Department of Geography of the University of Minnesota from ERTS photos. A line printer grey map of band 7 was used to register the ERTS data with lakes and rivers on the ground truth map. The 4/24/69 aerial photo for the ground truth map was taken by Mark Hurd Aerial Surveys, Inc. The University of Minnesota land use map was the principal source of ground truth with the aerial photo being used as a verification of the uses assigned on that map.

Land use, for this particular Iron Range study, was divided into nine classes:

Class	Name	Number
1	Hardwood	11
2	Conifer	12
3	Water	31
4	Water in mines	32
5	Wetlands	40
6	Mines	61
7	Tailings	62
8	Open	70
9	Urban	80

These classes represented all the distinct land uses within the area. When a class sample was identified on the density level map, it was delineated and labeled. This procedure was continued until the number of samples within each class was sufficient to establish the class' character. With the character of each class determined, automatic classification could proceed.

In an effort to decrease the number of classes, the original nine classes were compressed to six in the following manner. The compression was done by using the minimum, maximum, mean, and standard deviation information as well as information about the land uses themselves. The nine classes were merged into six as follows:

Feature	New Class
Hardwood (11)	1
Conifer (12)	2
Water (31)	3
Water in mines (32)	3
Wetlands (40)	4
Mines (61)	4
Tailings (62)	4
Open (70)	5
Urban (80)	6

Wetlands were put with mines and tailings not only because their numerical values were similar; but, also, because their occurrence in this area seems to be mostly the result of mines and tailings (i.e., part of tailing ponds).

The training set contained a total of 2855 samples with 543 in class 1, 381 in class 2, 327 in class 3, 1256 in class 4, 214 in class 5, and 134 in class 6. In order to determine the separability of classes, the mean, stan-

dard deviation, minimum and maximum were computed for each class, for each spectral band and for each of the six channels per band.

The difference between the means for each class compared to the difference between means per channel is a measure of signal to noise ratio. This data, which is listed in detail in Table II, is plotted in Figure 2. The ratio of σ^2 (class)/σ^2 (channel) indicates a relative signal to noise figure. If this ratio decreased to unity, it would indicate that the banding effect overshadowed class differences and a need to normalize the channels.

The training data used to train the automatic classifier; that is, a set of weights, were determined such that the training set was clustered around points equidistant from one another. When this set of weights is determined, an iteration is performed on the a priori probabilities to reduce the number of misses.

Each iteration of the K-Class program provides the number of misses along with a confusion matrix by percent and by samples. The minimum number of errors was obtained in the sixth iteration. An overall accuracy of 77.6% was achieved (see Table III).

Upon reaching the minimum number of misclassifications, the iterations are terminated and a list of the misses at that point are printed out. The list contains the ID number of the element, its assigned class, and the class decided upon by the K-Class program. A comparison of this information with the input data from the four bands and the minimum-maximum data for each class aided in the explanation of particular misses.

Some misses arise from the erroneous designation of training samples. A comparison of the printout of sample location, designation, and data of each band and the maximum-minimum information with the density level map reveals

Table II.

Probability Distribution of the Data as a Function of Channel Over the Four Spectral Bands

```
------ CLASS :  11 ------   1 -- Hardwood

CHANNEL =  1
    BAND. = 4, MIN.= 21, MAX.= 31, MEAN = 25.236, ST.DV. = 1.7739
    BAND. = 5, MIN.= 17, MAX.= 28, MEAN = 24.371, ST.DV. = 2.7371
    BAND. = 6, MIN.= 27, MAX.= 38, MEAN = 31.787, ST.DV. = 2.2609
```

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

```
    BAND. = 7, MIN.= 15, MAX.= 23, MEAN = 18.281, ST.DV. = 1.6421

CHANNEL =  2
    BAND. = 4, MIN.= 20, MAX.= 28, MEAN = 24.812, ST.DV. = 1.5337
    BAND. = 5, MIN.= 17, MAX.= 28, MEAN = 24.150, ST.DV. = 2.7391
    BAND. = 6, MIN.= 27, MAX.= 39, MEAN = 32.750, ST.DV. = 2.4213
    BAND. = 7, MIN.= 16, MAX.= 23, MEAN = 18.337, ST.DV. = 1.5325

CHANNEL =  3
    BAND. = 4, MIN.= 22, MAX.= 28, MEAN = 24.551, ST.DV. = 1.9488
    BAND. = 5, MIN.= 16, MAX.= 27, MEAN = 23.955, ST.DV. = 2.5743
    BAND. = 6, MIN.= 26, MAX.= 39, MEAN = 32.135, ST.DV. = 2.8490
    BAND. = 7, MIN.= 14, MAX.= 24, MEAN = 18.225, ST.DV. = 1.7848

CHANNEL =  4
    BAND. = 4, MIN.= 21, MAX.= 28, MEAN = 24.615, ST.DV. = 1.8219
    BAND. = 5, MIN.= 17, MAX.= 30, MEAN = 24.239, ST.DV. = 2.3653
    BAND. = 6, MIN.= 25, MAX.= 40, MEAN = 31.725, ST.DV. = 2.9858
    BAND. = 7, MIN.= 15, MAX.= 25, MEAN = 18.550, ST.DV. = 1.7995

CHANNEL =  5
    BAND. = 4, MIN.= 21, MAX.= 27, MEAN = 23.776, ST.DV. = 1.3630
    BAND. = 5, MIN.= 16, MAX.= 27, MEAN = 23.855, ST.DV. = 2.6985
    BAND. = 6, MIN.= 26, MAX.= 40, MEAN = 33.395, ST.DV. = 2.27[illegible]
    BAND. = 7, MIN.= 16, MAX.= 24, MEAN = 18.395, ST.DV. = 1.6629

CHANNEL =  6
    BAND. = 4, MIN.= 21, MAX.= 28, MEAN = 24.720, ST.DV. = 1.8925
    BAND. = 5, MIN.= 20, MAX.= 32, MEAN = 24.820, ST.DV. = 2.3511
    BAND. = 6, MIN.= 28, MAX.= 41, MEAN = 32.370, ST.DV. = 2.8448
    BAND. = 7, MIN.= 15, MAX.= 23, MEAN = 18.450, ST.DV. = 1.7400
```

Table II (continued)

```
----- CLASS = 12 -----        2 -- Conifer

CHANNEL = 1                              flower-top
    BAND* = 4, MIN.= 17, MAX.= 24, MEAN = 20.800, ST.DV. = 1.7310
    BAND* = 5, MIN.= 11, MAX.= 22, MEAN = 16.491, ST.DV. = 2.7953
    BAND* = 6, MIN.= 17, MAX.= 27, MEAN = 23.200, ST.DV. = 2.3618
    BAND* = 7, MIN.=  9, MAX.= 15, MEAN = 12.964, ST.DV. = 1.1113

CHANNEL = 2
    BAND* = 4, MIN.= 17, MAX.= 25, MEAN = 21.038, ST.DV. = 1.7863
    BAND* = 5, MIN.= 13, MAX.= 22, MEAN = 16.718, ST.DV. = 2.5110
    BAND* = 6, MIN.= 17, MAX.= 29, MEAN = 22.974, ST.DV. = 2.4754
    BAND* = 7, MIN.=  9, MAX.= 15, MEAN = 12.179, ST.DV. = 1.4566

CHANNEL = 3
    BAND* = 4, MIN.= 18, MAX.= 23, MEAN = 20.302, ST.DV. = 1.3474
    BAND* = 5, MIN.= 12, MAX.= 21, MEAN = 15.442, ST.DV. = 2.4993
    BAND* = 6, MIN.= 17, MAX.= 27, MEAN = 21.895, ST.DV. = 2.7578
    BAND* = 7, MIN.=  9, MAX.= 16, MEAN = 11.942, ST.DV. = 1.4089

CHANNEL = 4
    BAND* = 4, MIN.= 17, MAX.= 23, MEAN = 19.754, ST.DV. = 1.3341
    BAND* = 5, MIN.= 11, MAX.= 22, MEAN = 15.217, ST.DV. = 2.4369
    BAND* = 6, MIN.= 17, MAX.= 26, MEAN = 21.623, ST.DV. = 2.1810
    BAND* = 7, MIN.=  9, MAX.= 14, MEAN = 12.232, ST.DV. = 1.2982

CHANNEL = 5
    BAND* = 4, MIN.= 17, MAX.= 23, MEAN = 19.367, ST.DV. = 1.5869
    BAND* = 5, MIN.= 11, MAX.= 21, MEAN = 15.184, ST.DV. = 2.7156


    BAND* = 6, MIN.= 15, MAX.= 25, MEAN = 22.041, ST.DV. = 2.5149
    BAND* = 7, MIN.=  8, MAX.= 14, MEAN = 11.306, ST.DV. = 1.2810

CHANNEL = 6
    BAND* = 4, MIN.= 18, MAX.= 25, MEAN = 20.727, ST.DV. = 1.8136
    BAND* = 5, MIN.= 11, MAX.= 26, MEAN = 16.341, ST.DV. = 3.5223
    BAND* = 6, MIN.= 17, MAX.= 29, MEAN = 22.773, ST.DV. = 2.5661
    BAND* = 7, MIN.=  8, MAX.= 16, MEAN = 12.295, ST.DV. = 1.6178
```

Table II (continued)

------ CLASS = 31 ------ 3 -- Water

```
CHANNEL =  1
    BAND* = 4, MIN.= 16, MAX.= 20, MEAN = 18.471, ST.DV. = 1.2184
    BAND* = 5, MIN.=  9, MAX.= 15, MEAN = 10.794, ST.DV. = 1.3456
    BAND* = 6, MIN.=  6, MAX.= 14, MEAN =  8.882, ST.DV. = 1.6227
    BAND* = 7, MIN.=  1, MAX.=  6, MEAN =  2.853, ST.DV. = 1.1149

CHANNEL =  2
    BAND* = 4, MIN.= 16, MAX.= 20, MEAN = 18.682, ST.DV. = 1.1633
    BAND* = 5, MIN.=  9, MAX.= 13, MEAN = 10.932, ST.DV. = 1.2685
    BAND* = 6, MIN.=  6, MAX.= 13, MEAN =  9.136, ST.DV. = 1.6038
    BAND* = 7, MIN.=  1, MAX.=  5, MEAN =  2.182, ST.DV. =  .9599

CHANNEL =  3
    BAND* = 4, MIN.= 16, MAX.= 20, MEAN = 18.268, ST.DV. =  .9113
    BAND* = 5, MIN.=  9, MAX.= 13, MEAN = 10.561, ST.DV. = 1.3261
    BAND* = 6, MIN.=  4, MAX.= 11, MEAN =  7.049, ST.DV. = 1.7384
    BAND* = 7, MIN.=  1, MAX.=  5, MEAN =  2.073, ST.DV. =  .8665

CHANNEL =  4
    BAND* = 4, MIN.= 17, MAX.= 21, MEAN = 18.091, ST.DV. = 1.1642
    BAND* = 5, MIN.=  8, MAX.= 15, MEAN = 10.485, ST.DV. = 1.7429
    BAND* = 6, MIN.=  4, MAX.= 12, MEAN =  7.212, ST.DV. = 2.0999
    BAND* = 7, MIN.=  1, MAX.=  5, MEAN =  2.424, ST.DV. = 1.1018

CHANNEL =  5
    BAND* = 4, MIN.= 15, MAX.= 21, MEAN = 17.400, ST.DV. = 2.1541
    BAND* = 5, MIN.=  9, MAX.= 14, MEAN = 11.100, ST.DV. = 1.9723
    BAND* = 6, MIN.=  6, MAX.= 13, MEAN =  8.400, ST.DV. = 2.6533
    BAND* = 7, MIN.=  1, MAX.=  5, MEAN =  2.500, ST.DV. = 1.2845

CHANNEL =  6
    BAND* = 4, MIN.= 17, MAX.= 22, MEAN = 18.864, ST.DV. = 1.6038
    BAND* = 5, MIN.=  9, MAX.= 15, MEAN = 11.091, ST.DV. = 1.4431
    BAND* = 6, MIN.=  7, MAX.= 13, MEAN =  9.227, ST.DV. = 1.3570
    BAND* = 7, MIN.=  2, MAX.=  5, MEAN =  3.091, ST.DV. = 1.1245
```


REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

Table II (continued)

----- CLASS = 32 ----- 4 -- Water in Mines

```
CHANNEL =  1
     BAND* = 4, MIN.= 19, MAX.= 24, MEAN = 21.241, ST.DV. = 1.8318
     BAND* = 5, MIN.= 10, MAX.= 17, MEAN = 14.276, ST.DV. = 1.9982
     BAND* = 6, MIN.=  7, MAX.= 14, MEAN = 10.759, ST.DV. = 1.6745
     BAND* = 7, MIN.=  2, MAX.=  6, MEAN =  3.621, ST.DV. =  .9970

CHANNEL =  2
     BAND* = 4, MIN.= 20, MAX.= 25, MEAN = 21.958, ST.DV. = 1.9035
     BAND* = 5, MIN.= 11, MAX.= 20, MEAN = 14.792, ST.DV. = 2.1981
     BAND* = 6, MIN.=  7, MAX.= 16, MEAN = 11.292, ST.DV. = 2.2818
     BAND* = 7, MIN.=  1, MAX.=  6, MEAN =  3.292, ST.DV. = 1.4855

CHANNEL =  3
     BAND* = 4, MIN.= 19, MAX.= 26, MEAN = 21.409, ST.DV. = 1.7751
     BAND* = 5, MIN.= 10, MAX.= 19, MEAN = 14.727, ST.DV. = 2.2600
     BAND* = 6, MIN.=  6, MAX.= 15, MEAN = 10.409, ST.DV. = 2.1247
     BAND* = 7, MIN.=  1, MAX.=  6, MEAN =  3.409, ST.DV. = 1.1544

CHANNEL =  4
     BAND* = 4, MIN.= 20, MAX.= 25, MEAN = 21.417, ST.DV. = 1.3202
     BAND* = 5, MIN.= 11, MAX.= 17, MEAN = 14.917, ST.DV. = 1.3202
     BAND* = 6, MIN.=  7, MAX.= 12, MEAN =  9.917, ST.DV. = 1.3202
     BAND* = 7, MIN.=  2, MAX.=  5, MEAN =  3.750, ST.DV. =  .8292

CHANNEL =  5
     BAND* = 4, MIN.= 16, MAX.= 23, MEAN = 20.947, ST.DV. = 1.7614
     BAND* = 5, MIN.= 11, MAX.= 16, MEAN = 14.105, ST.DV. = 1.4103
     BAND* = 6, MIN.=  8, MAX.= 12, MEAN =  9.526, ST.DV. = 1.0939
     BAND* = 7, MIN.=  2, MAX.=  3, MEAN =  2.684, ST.DV. =  .4648

CHANNEL =  6
     BAND* = 4, MIN.= 16, MAX.= 25, MEAN = 21.040, ST.DV. = 1.3705
     BAND* = 5, MIN.= 11, MAX.= 18, MEAN = 14.520, ST.DV. = 2.0024
     BAND* = 6, MIN.=  8, MAX.= 15, MEAN = 10.760, ST.DV. = 1.5819
     BAND* = 7, MIN.=  2, MAX.=  6, MEAN =  3.360, ST.DV. =  .9749
```

Table II (continued)

----- CLASS = 40 ----- 5 -- Wetlands (w/Tailing & Mines)

noise bottom

CHANNEL	BAND	MIN.	MAX.	MEAN	ST.DV.
1	4	19	27	23.214	2.4836
	5	14	27	19.333	3.1522
	6	17	29	23.524	3.1866
	7	6	15	12.310	1.8960
2	4	19	25	22.846	2.1428
	5	13	23	18.949	3.0963
	6	12	29	23.103	3.3034
	7	4	16	11.538	2.2854
3	4	18	33	22.632	2.7855
	5	14	30	19.395	3.3837
	6	18	29	23.105	2.9539
	7	8	15	12.105	1.7739
4	4	20	25	23.024	1.6303
	5	14	23	19.244	2.2172
	6	17	28	23.317	2.8149
	7	5	16	12.634	2.0572
5	4	18	27	22.882	1.7784
	5	15	25	19.353	1.9385
	6	14	30	24.176	4.1761
	7	4	16	11.676	2.6977
6	4	20	27	23.730	1.9952
	5	15	29	21.027	2.9727
	6	13	29	23.568	3.7166
	7	6	15	11.784	2.2317

Table II (continued)

```
----- CLASS = 61 -----    6 -- Mines

CHANNEL =  1          noise bottom
     BAND* = 4, MIN.= 21, MAX.= 28, MEAN = 24.307, ST.DV. = 1.7037
     BAND* = 5, MIN.= 15, MAX.= 30, MEAN = 23.048, ST.DV. = 2.9285
     BAND* = 6, MIN.= 12, MAX.= 29, MEAN = 23.000, ST.DV. = 3.7352
     BAND* = 7, MIN.=  5, MAX.= 16, MEAN = 10.157, ST.DV. = 2.1368

CHANNEL =  2
     BAND* = 4, MIN.= 20, MAX.= 28, MEAN = 24.922, ST.DV. = 1.4924
     BAND* = 5, MIN.= 14, MAX.= 28, MEAN = 23.189, ST.DV. = 2.9511
     BAND* = 6, MIN.= 11, MAX.= 32, MEAN = 24.156, ST.DV. = 3.8899
     BAND* = 7, MIN.=  4, MAX.= 16, MEAN =  9.844, ST.DV. = 2.5029

CHANNEL =  3
     BAND* = 4, MIN.= 20, MAX.= 29, MEAN = 24.530, ST.DV. = 1.9096
     BAND* = 5, MIN.= 19, MAX.= 30, MEAN = 23.361, ST.DV. = 2.4768
     BAND* = 6, MIN.= 15, MAX.= 31, MEAN = 23.169, ST.DV. = 3.2032
     BAND* = 7, MIN.=  6, MAX.= 16, MEAN =  9.482, ST.DV. = 1.8971

CHANNEL =  4
     BAND* = 4, MIN.= 21, MAX.= 28, MEAN = 24.212, ST.DV. = 1.5305
     BAND* = 5, MIN.= 14, MAX.= 27, MEAN = 22.787, ST.DV. = 2.5870
     BAND* = 6, MIN.= 14, MAX.= 30, MEAN = 23.200, ST.DV. = 3.3667
     BAND* = 7, MIN.=  5, MAX.= 16, MEAN = 10.000, ST.DV. = 2.3611

CHANNEL =  5
     BAND* = 4, MIN.= 21, MAX.= 32, MEAN = 23.772, ST.DV. = 1.4073
     BAND* = 5, MIN.= 14, MAX.= 33, MEAN = 22.174, ST.DV. = 2.8876
     BAND* = 6, MIN.= 13, MAX.= 34, MEAN = 23.674, ST.DV. = 3.4801
     BAND* = 7, MIN.=  4, MAX.= 15, MEAN =  9.620, ST.DV. = 2.0999

CHANNEL =  6
     BAND* = 4, MIN.= 22, MAX.= 32, MEAN = 24.851, ST.DV. = 1.8725
     BAND* = 5, MIN.= 18, MAX.= 32, MEAN = 23.103, ST.DV. = 2.5910
     BAND* = 6, MIN.= 13, MAX.= 36, MEAN = 23.713, ST.DV. = 3.7475
     BAND* = 7, MIN.=  5, MAX.= 17, MEAN = 10.368, ST.DV. = 2.3445
```

Table II (continued)

```
------ CLASS = 62 ------        7 -- Tailings
                    noise bottom
CHANNEL =  1
     BAND. = 4, MIN.= 20, MAX.= 32, MEAN = 24.830, ST.DV. = 2.5832
     BAND. = 5, MIN.= 16, MAX.= 35, MEAN = 23.691, ST.DV. = 4.8664
     BAND. = 6, MIN.= 18, MAX.= 34, MEAN = 25.096, ST.DV. = 3.8039
     BAND. = 7, MIN.=  7, MAX.= 15, MEAN = 11.394, ST.DV. = 1.9198

CHANNEL =  2
     BAND. = 4, MIN.= 20, MAX.= 31, MEAN = 24.539, ST.DV. = 2.3178
     BAND. = 5, MIN.= 16, MAX.= 34, MEAN = 22.551, ST.DV. = 4.3005
     BAND. = 6, MIN.= 17, MAX.= 35, MEAN = 24.640, ST.DV. = 3.5450
     BAND. = 7, MIN.=  6, MAX.= 14, MEAN = 10.787, ST.DV. = 2.1799

CHANNEL =  3
     BAND. = 4, MIN.= 20, MAX.= 39, MEAN = 23.845, ST.DV. = 3.2385
     BAND. = 5, MIN.= 15, MAX.= 34, MEAN = 21.976, ST.DV. = 3.7606
     BAND. = 6, MIN.= 17, MAX.= 31, MEAN = 23.869, ST.DV. = 2.9672
     BAND. = 7, MIN.=  6, MAX.= 16, MEAN = 10.905, ST.DV. = 2.3331

CHANNEL =  4
     BAND. = 4, MIN.= 20, MAX.= 37, MEAN = 24.465, ST.DV. = 3.0220
     BAND. = 5, MIN.= 15, MAX.= 38, MEAN = 23.209, ST.DV. = 5.2697
     BAND. = 6, MIN.= 17, MAX.= 38, MEAN = 25.035, ST.DV. = 4.1469
     BAND. = 7, MIN.=  5, MAX.= 16, MEAN = 11.453, ST.DV. = 2.2447

CHANNEL =  5
     BAND. = 4, MIN.= 19, MAX.= 29, MEAN = 23.207, ST.DV. = 1.8329
     BAND. = 5, MIN.= 16, MAX.= 33, MEAN = 21.793, ST.DV. = 4.0004
     BAND. = 6, MIN.= 14, MAX.= 34, MEAN = 24.770, ST.DV. = 4.0278
     BAND. = 7, MIN.=  5, MAX.= 15, MEAN = 10.770, ST.DV. = 2.0550

CHANNEL =  6
     BAND. = 4, MIN.= 20, MAX.= 32, MEAN = 24.900, ST.DV. = 2.6140
     BAND. = 5, MIN.= 15, MAX.= 35, MEAN = 23.971, ST.DV. = 4.8843
     BAND. = 6, MIN.= 15, MAX.= 35, MEAN = 25.386, ST.DV. = 3.9939
     BAND. = 7, MIN.=  6, MAX.= 16, MEAN = 11.429, ST.DV. = 2.0603
```

Table II (continued)

```
------ CLASS = 70 ------        8 -- Open

CHANNEL =  1
     BAND. = 4, MIN.= 20, MAX.= 27, MEAN = 24.237, ST.DV. = 1.8272
     BAND. = 5, MIN.= 14, MAX.= 27, MEAN = 18.158, ST.DV. = 2.6904
     BAND. = 6, MIN.= 29, MAX.= 45, MEAN = 37.605, ST.DV. = 3.9439
     BAND. = 7, MIN.= 16, MAX.= 31, MEAN = 23.526, ST.DV. = 3.0927

CHANNEL =  2
     BAND. = 4, MIN.= 20, MAX.= 28, MEAN = 24.367, ST.DV. = 1.7693
     BAND. = 5, MIN.= 14, MAX.= 26, MEAN = 18.959, ST.DV. = 3.0968
     BAND. = 6, MIN.= 29, MAX.= 46, MEAN = 37.633, ST.DV. = 4.6585
     BAND. = 7, MIN.= 17, MAX.= 30, MEAN = 22.755, ST.DV. = 3.0741

CHANNEL =  3
     BAND. = 4, MIN.= 22, MAX.= 28, MEAN = 23.680, ST.DV. = 1.7826
     BAND. = 5, MIN.= 14, MAX.= 23, MEAN = 18.100, ST.DV. = 2.2204
     BAND. = 6, MIN.= 31, MAX.= 46, MEAN = 38.080, ST.DV. = 4.0440
     BAND. = 7, MIN.= 17, MAX.= 28, MEAN = 22.840, ST.DV. = 2.7376

CHANNEL =  4
     BAND. = 4, MIN.= 21, MAX.= 27, MEAN = 23.839, ST.DV. = 1.8854
     BAND. = 5, MIN.= 15, MAX.= 26, MEAN = 18.548, ST.DV. = 3.3199
     BAND. = 6, MIN.= 28, MAX.= 44, MEAN = 37.645, ST.DV. = 4.6458
     BAND. = 7, MIN.= 14, MAX.= 27, MEAN = 22.371, ST.DV. = 3.2897

CHANNEL =  5
     BAND. = 4, MIN.= 19, MAX.= 25, MEAN = 23.267, ST.DV. = 1.5261
     BAND. = 5, MIN.= 16, MAX.= 26, MEAN = 18.933, ST.DV. = 2.5421
     BAND. = 6, MIN.= 30, MAX.= 50, MEAN = 38.467, ST.DV. = 4.9648
     BAND. = 7, MIN.= 19, MAX.= 26, MEAN = 22.533, ST.DV. = 2.1250

CHANNEL =  6
     BAND. = 4, MIN.= 21, MAX.= 27, MEAN = 23.645, ST.DV. = 2.0407
     BAND. = 5, MIN.= 13, MAX.= 26, MEAN = 18.065, ST.DV. = 2.8163
     BAND. = 6, MIN.= 28, MAX.= 49, MEAN = 38.806, ST.DV. = 4.8952
     BAND. = 7, MIN.= 17, MAX.= 30, MEAN = 23.290, ST.DV. = 3.1335
```

Table II (concluded)

```
----- CLASS = 80 -----     9 -- Urban

CHANNEL =  1
    BAND* = 4, MIN.= 24, MAX.= 31, MEAN = 26.970, ST.DV. = 1.3814
    BAND* = 5, MIN.= 18, MAX.= 27, MEAN = 22.727, ST.DV. = 1.5230
    BAND* = 6, MIN.= 29, MAX.= 34, MEAN = 31.091, ST.DV. = 1.5048
    BAND* = 7, MIN.= 13, MAX.= 19, MEAN = 16.515, ST.DV. = 1.4590

CHANNEL =  2
    BAND* = 4, MIN.= 24, MAX.= 28, MEAN = 26.320, ST.DV. = 1.2238
    BAND* = 5, MIN.= 20, MAX.= 26, MEAN = 22.080, ST.DV. = 1.5728
    BAND* = 6, MIN.= 27, MAX.= 39, MEAN = 31.080, ST.DV. = 2.7701
    BAND* = 7, MIN.= 14, MAX.= 22, MEAN = 16.320, ST.DV. = 1.7600

CHANNEL =  3
    BAND* = 4, MIN.= 26, MAX.= 33, MEAN = 28.000, ST.DV. = 1.9149
    BAND* = 5, MIN.= 20, MAX.= 27, MEAN = 23.333, ST.DV. = 2.4944
    BAND* = 6, MIN.= 26, MAX.= 35, MEAN = 31.750, ST.DV. = 2.9190
    BAND* = 7, MIN.= 13, MAX.= 20, MEAN = 16.583, ST.DV. = 2.2898

CHANNEL =  4
    BAND* = 4, MIN.= 27, MAX.= 32, MEAN = 28.000, ST.DV. = 1.4606
    BAND* = 5, MIN.= 22, MAX.= 30, MEAN = 24.400, ST.DV. = 1.9933
    BAND* = 6, MIN.= 26, MAX.= 38, MEAN = 33.133, ST.DV. = 3.5565
    BAND* = 7, MIN.= 13, MAX.= 20, MEAN = 17.600, ST.DV. = 2.0591

CHANNEL =  5
    BAND* = 4, MIN.= 25, MAX.= 29, MEAN = 27.000, ST.DV. = 1.2978
    BAND* = 5, MIN.= 21, MAX.= 26, MEAN = 23.632, ST.DV. = 1.9252
    BAND* = 6, MIN.= 27, MAX.= 38, MEAN = 33.316, ST.DV. = 3.2453
    BAND* = 7, MIN.= 11, MAX.= 20, MEAN = 16.421, ST.DV. = 2.3241

CHANNEL =  6
    BAND* = 4, MIN.= 25, MAX.= 32, MEAN = 27.700, ST.DV. = 1.3204
    BAND* = 5, MIN.= 22, MAX.= 29, MEAN = 24.533, ST.DV. = 2.0450
    BAND* = 6, MIN.= 26, MAX.= 36, MEAN = 31.533, ST.DV. = 3.4325
    BAND* = 7, MIN.= 11, MAX.= 21, MEAN = 16.167, ST.DV. = 2.5959

*STOP* 00000000
```

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

1075-16312

Figure 4.

Class and Channel Dispersion

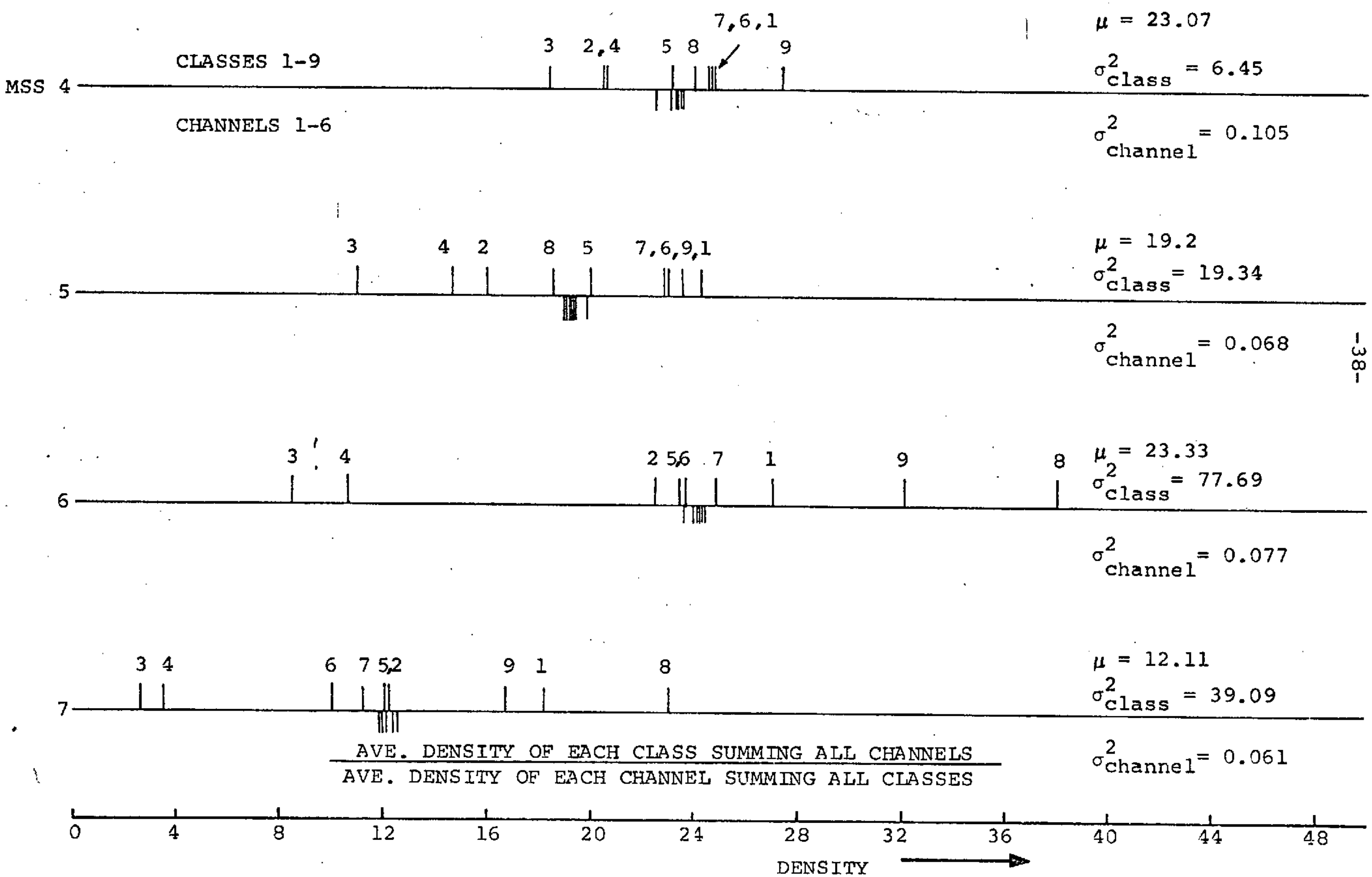


Table III.

Confusion Matrix for the Trout Lake Area

Confusion Matrix in Percent:

Ground Truth		Classifier Assignment 1/	2/	3/	4/	5/	6/
Hardwood	1/	77•3	•4	•0	4•6	9•6	8•1
Conifer	2/	10•0	80•3	2•1	5•8	•3	1•6
Water	3/	•0	3•1	95•4	1•5	•0	•0
Mines	4/	8•6	9•6	4•2	70•1	•3	7•2
Open	5/	7•5	•5	•0	•0	87•9	4•2
Urban	6/	5•2	•0	•0	11•9	3•0	79•9

Confusion Matrix in Samples:

420	2	0	25	52	44
38	306	8	22	1	6
0	10	312	5	0	0
108	121	53	880	4	90
16	1	0	0	188	9
7	0	0	16	4	107

Overall accuracy: $\frac{2855-642}{2855} = 77.6\%$

many of these erroneous designations. Often the values within a data block were very uniform except for one or two points on the periphery. These peripheral points usually fell on or very near the minimum-maximum sheets. Upon checking the location of these points on the density level map, it was, in most cases, found that they were not actually part of the class with which they had been grouped, but were part of a neighboring class.

The inclusion of these peripheral points in the wrong class caused an alteration of the minimum-maximum values of that class. That is to say, the inclusion of some hardwoods with open would cause the distinctions between the two classes, expressed by the minimum-maximum values for each channel to be blurred. Thus, numerous occasions would arise when an element could be classified in one of two ways. A correction of these erroneous training samples would eliminate some of the extremes in classes (i.e., lower the maximum and raise the minimum); therefore, further separating the classes and allowing the K-Class program to operate more effectively.

In two instances, the wrong column or row number was punched onto the computer cards. Thus a whole training set data block was erroneously classed. Correction of this error would have the same effect as the correction of the previous type: elimination of extreme values in classes.

The misses thus far discussed have been human errors. Others occurred because of the overlap between classes. For two classes to be distinguished from one another 100 percent of the time, none of their values may overlap; i.e., the maximum of one must be less than the minimum of the other. This is generally not the case.

For example, the values for open and conifer overlap in bands 4 and 5 but not in 6 and 7. Conifer and urban overlap in bands 5 and 7, but not in 4 and 6. The two classes of water and conifer overlap in bands 4 and 5, but not in 6 and 7. Natural water can be distinguished from water in mines by band 4 only. This is an important distinction; because, natural waters (lakes, rivers, reservoirs) can be used for recreational purposes, while water in mines can not. Hence, when taking a total of all useable waters in an area, water in mines should not be included. Hardwoods can be distinguished from tailings, mines, and wetlands by band 7 only. Whereas hardwoods and conifers can be distinguished by 6 and 7, they cannot by 4 and 5. In over half of the misses that occurred the value of at least one of the bands did not fit within the minimum-maximum values for the class decision.

Some classes, however, overlap at least to some degree in all of the four bands. Mines and tailings overlap, but this is not of extreme importance because they are both in the class extractive. Hardwood and open completely overlap, as do hardwood and urban. In the case of the former, however, their means are different enough to suggest that the elimination of the erroneous peripheral data elements mentioned earlier may cause their values to separate. This seems much less likely in the

case of the latter. Open and urban also overlap on all four bands. Their overlaps are slight enough, especially in band 4, that they could at least be lessened by eliminating erroneous data.

A comparison was made of the data from the Trout Lake area with that of Pineville, located northwest on the same photo. The two areas were analyzed using a different method for deriving the ground truth and therefore the training samples. The various classes were compared using the minimum-maximum sheets of the two areas.

Values for hardwoods in the two areas matched in bands 6 and 7 with the mean for Pineville being slightly higher, but not in bands 4 and 5, where the Trout Lake means are higher. Conifers matched except for band 5 with, again, Trout Lake's mean being higher. Urban matched well in all four bands. Open matched in band 6, but only overlaps in bands 4, 5, and 7 with Trout Lake's mean consistently higher. No comparison could be made of extractive due to the differences in the extractive classes used (extractive 1, 2, 3 and 4 in the Pineville study verses mines, tailings, and water in mines in the Trout Lake study.)

These differences may have occurred due to the relative locations of the areas on the photo or more likely, they are due to differences in ground truth interpretation, since the two studies match in such easy to identify classes as urban.

The ground truth for Pineville was delineated with high altitude and regular aerial photos, while Trout Lake was delineated using landuse interpretations from ERTS photos with aerial photos as a backup. In the latter case, qualified land use interpreters had already decided what could or could not be designated from satellite photos.

Since ground truth for the two areas from photo 1075-16312 was not obtained in the same manner, the comparison of the results of the two areas cannot be considered conclusive. In the next experiment to be described, two areas (Pineville and Virginia) on the same photo were analyzed using the same procedure for obtaining the ground truth. This study was performed to determine the transferrability of class interval designations from one part of an ERTS image to another.

Pineville Virginia Area

The next ERTS image analyzed was 1057-16311, a September 18th coverage over Virginia and Pineville in northern Minnesota. These two cities are approximately 20 miles apart. In this experiment the training set was obtained over Pineville and the training weights obtained were applied to the Virginia area. This particular image was cloud-free but contained banding on MSS 4 and MSS 6.

One of the six channels on band 7 (channel 2) was attenuated with respect to the other five. This produced an image with repetitive bands spaced six scans apart. A signal of this type and its Fourier Transform is shown in Figure 5. The Fourier Transform computed from the ERTS data in which every sixth line is attenuated is shown in Figure 6. This spectrum was computed from word 200 of 256 scan lines; that is, from one column of the image

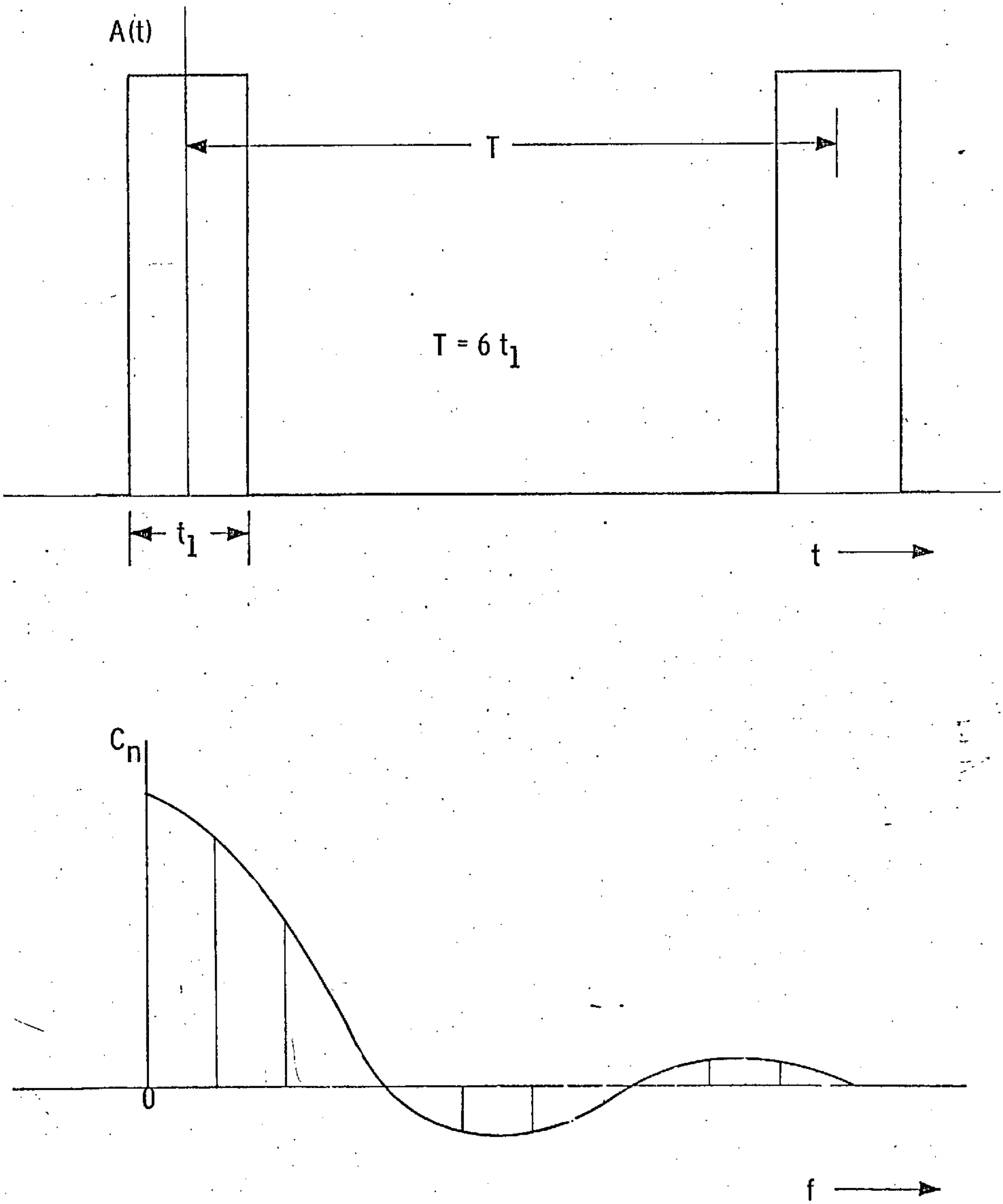
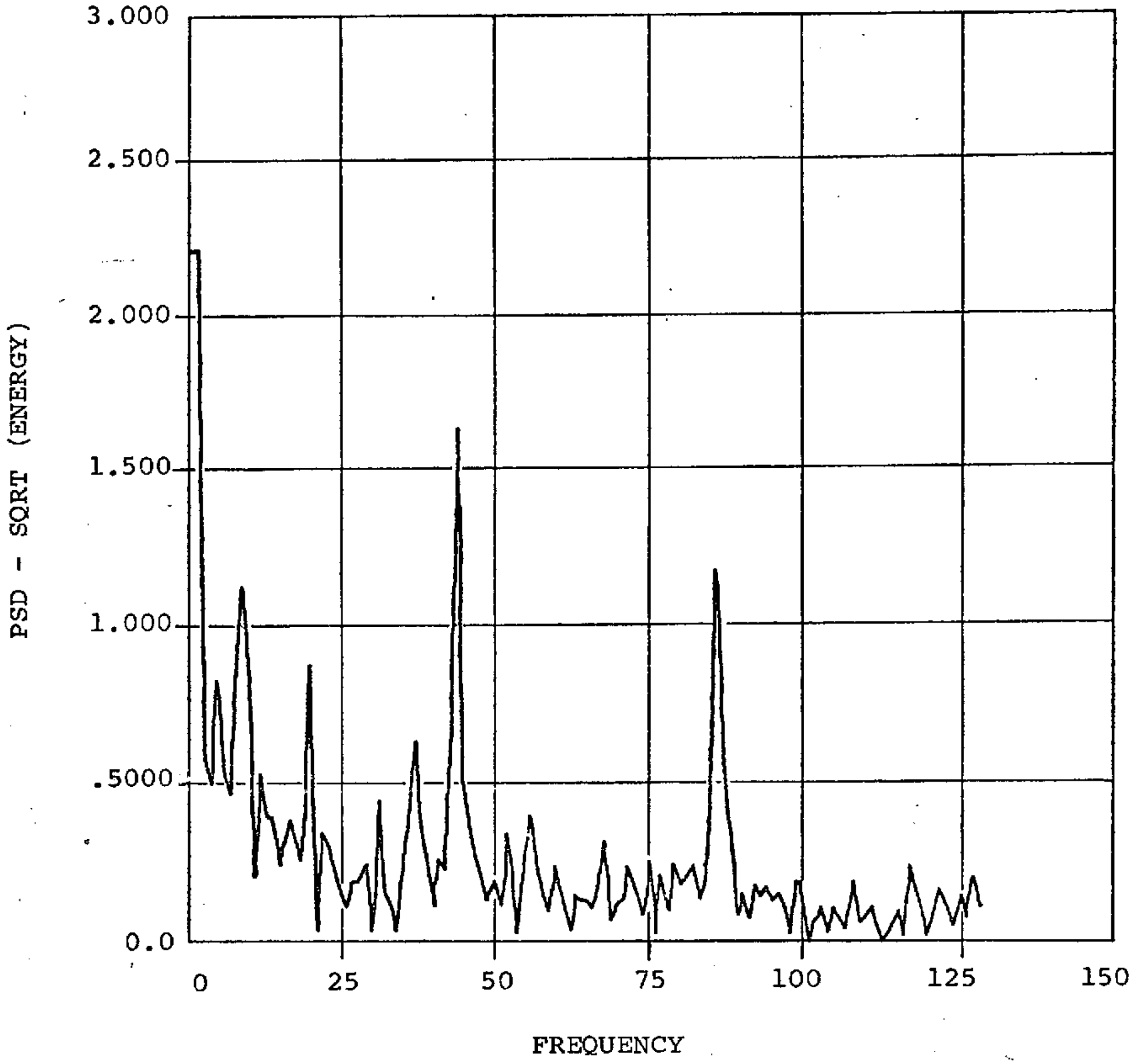


Figure 5.

Fourier-series Coefficients for a repeated Rectangular Pulse



PINEVILLE (1057-16311)
BAND 4

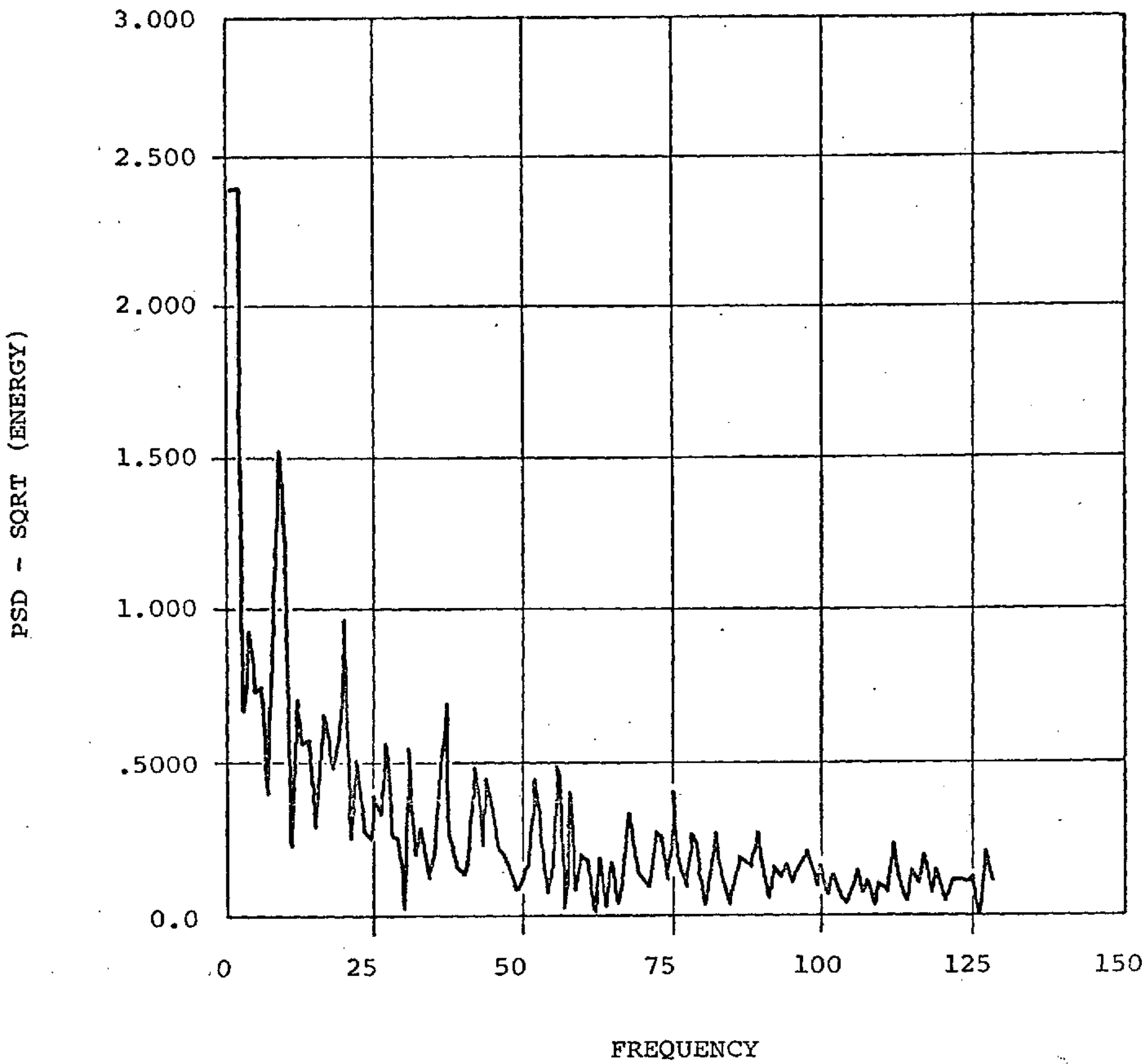
Figure 6. Power Spectral Density

matrix. A direct comparison can be made between the spectrum from band 4 in Figure 6 to the spectrum of band 5 on Figure 7. Both have the same vertical scale. Since band 5 does not have banding (all 6 channels have equal gain), the two spikes at 43 and 86 are absent.

Band 6 had one channel (channel 5) with intermittent drop-outs providing a broad spectrum of noise, as shown in Figure 8. This curve can be compared with band 7 in Figure 9, a band having no drop-outs. As expected, power spectrum on Figure 9 is much lower.

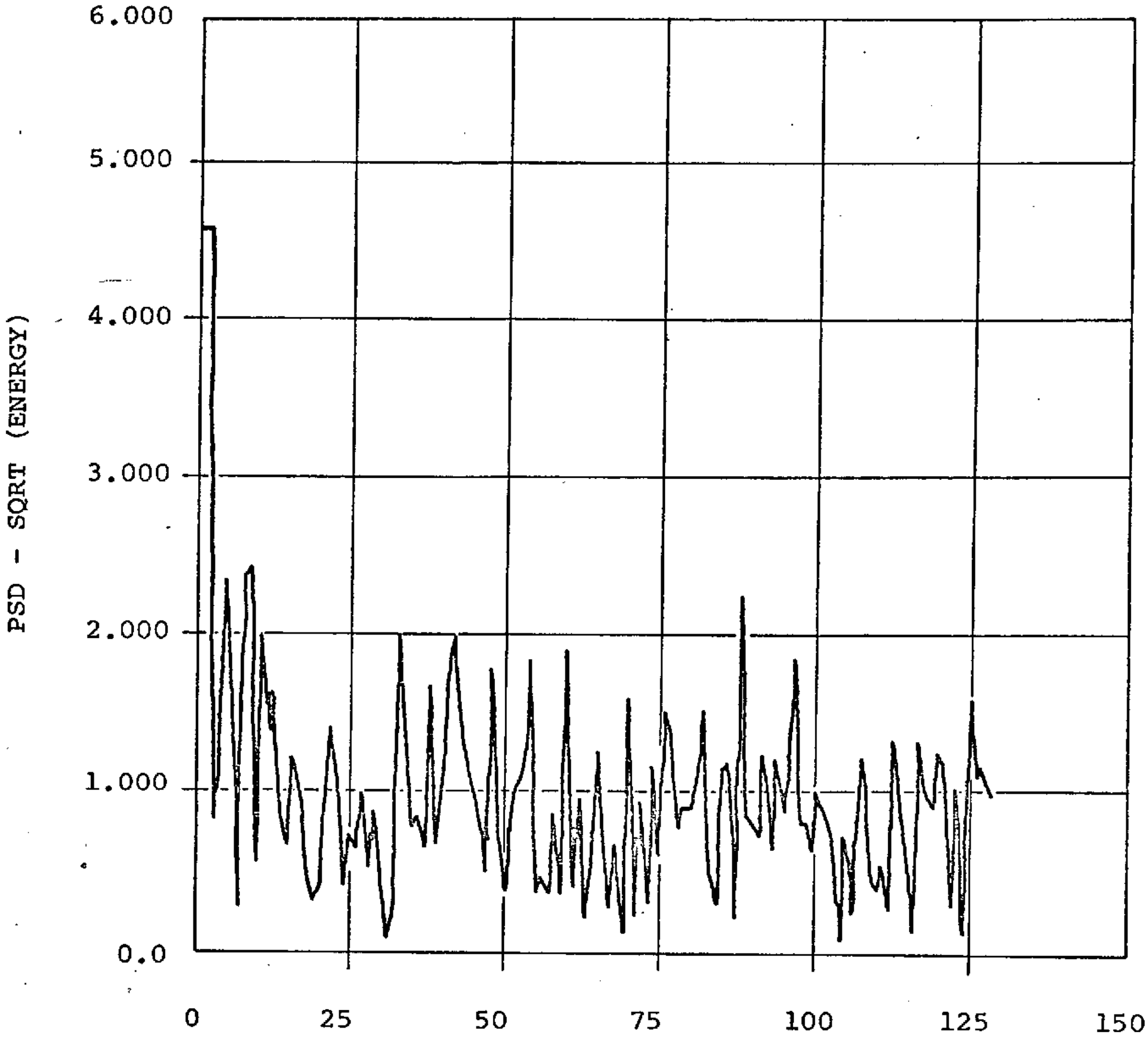
Because of the banding problem, three classifier training sets were derived; that is, classification weights were determined for channels having consistent gain 1, 3, 4, 6. Separate weights were obtained for channel 2 because of the attenuation on band 4. Separate weights were also obtained for channel 5 because of its sporadic dropout on band 6. Thus one set of weights were used on scan lines 1, 3, 4, 6, 7, 9, 10, 12, etc. Another set of weights were used on scan lines 2, 8, 14, 20, etc. A third set of weights were used for scan lines 5, 11, 17, 23, etc.

The classification accuracy for the three is approximately equal. The percentage confusion matrix for six classes based on four channels 1, 3, 4, 6 of data is shown in Table IV.



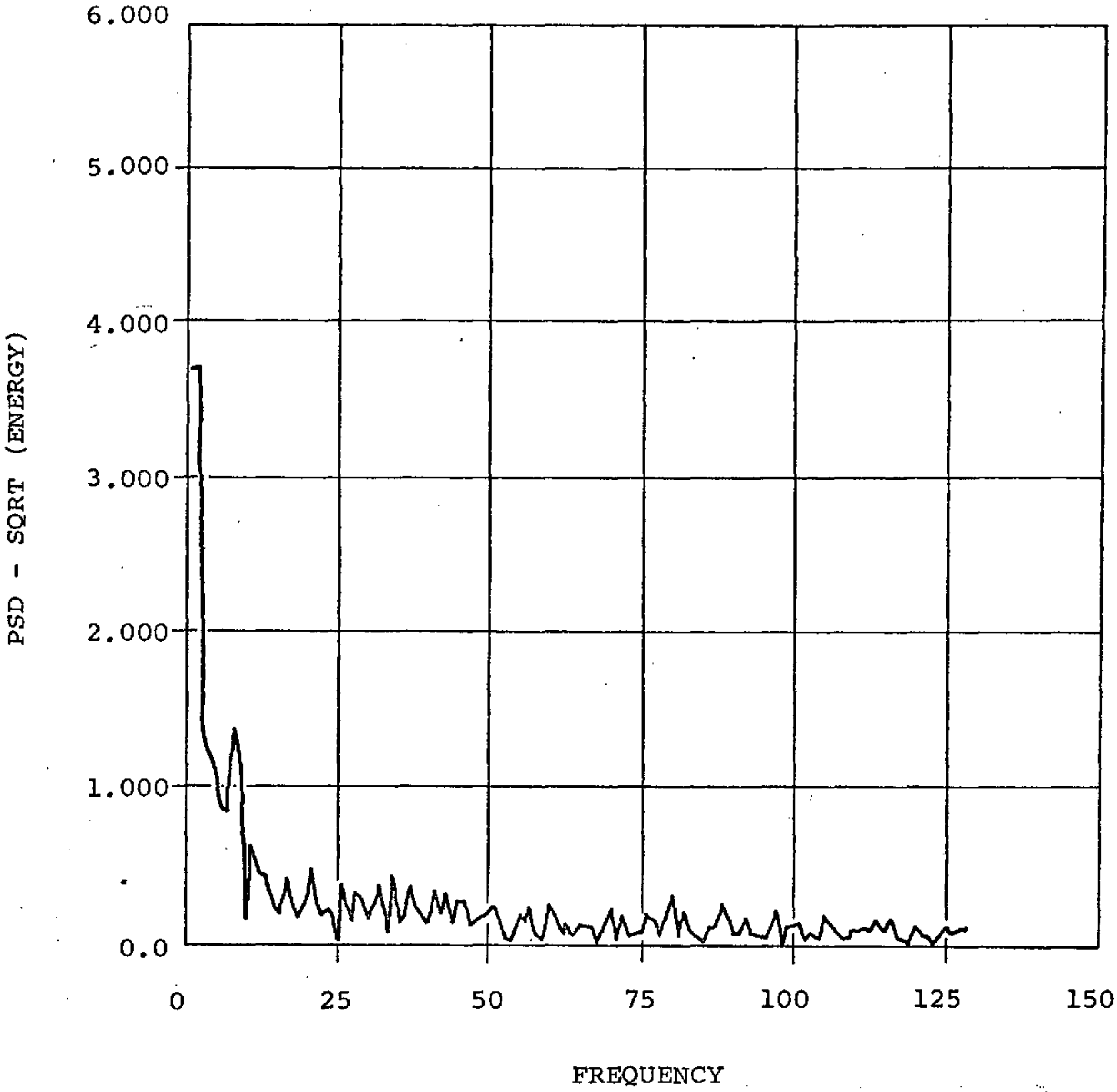
PINEVILLE (1057-16311)
BAND 5

Figure 7. Power Spectral Density



PINEVILLE (1057-16311)
BAND 6

Figure 8. Power Spectral Density



PINEVILLE (1057-16311)
BAND 7

Figure 9. Power Spectral Density

Table IV.

Pineville Confusion Matrix in Percent

Ground truth \ Classifier Assign.	Hdwd	Con	Res	Open	Mine	Tailings
Hardwood	89.9	4.3	.6	5.2	0	0
Conifer	3.0	89.7	0	.4	.4	6.5
Residential	0	0	95.7	3.2	1.1	0
Open	33.7	1.2	9.3	55.8	0	0
Open Pit Mine	0	.9	6.5	.9	89.8	1.9
Tailings	0	7.5	0	0	2.5	89.9

The overall accuracy for the data from these four channels was 87.5%. The overall accuracy is computed by summing the samples on the diagonal of the sample confusion matrix shown in Table V and dividing this sum by the total of all samples.

Table V.

Pineville Confusion Matrix, Samples

Ground Truth \ Classifier Assign.	H	C	R	O	M	T
Hardwood	295	14	2	17	0	0
Conifer	7	208	0	1	1	15
Residential	0	0	89	3	1	0
Open	29	1	8	48	0	0
Open Pit Mine	0	1	7	1	97	2
Tailings	0	12	0	0	4	143

The classification accuracy based on the data from channel 2 was 89.7%. The classification accuracy computed from channel 5 was 86.5%. One would expect this to be lowest because of the sporadic dropout. One would also expect that the accuracy from a single channel of data would be higher than from a combination of channels having slight variations in gain. This is evidenced by the performance on channel 2, where the accuracy is highest even though the signal is attenuated.

The thematic map generated for the Pineville area is shown in Figure 10. The classes in going from light to dark are Open, Conifer, Hardwood, Urban, Tailings, Open Pit Mines, and Water.

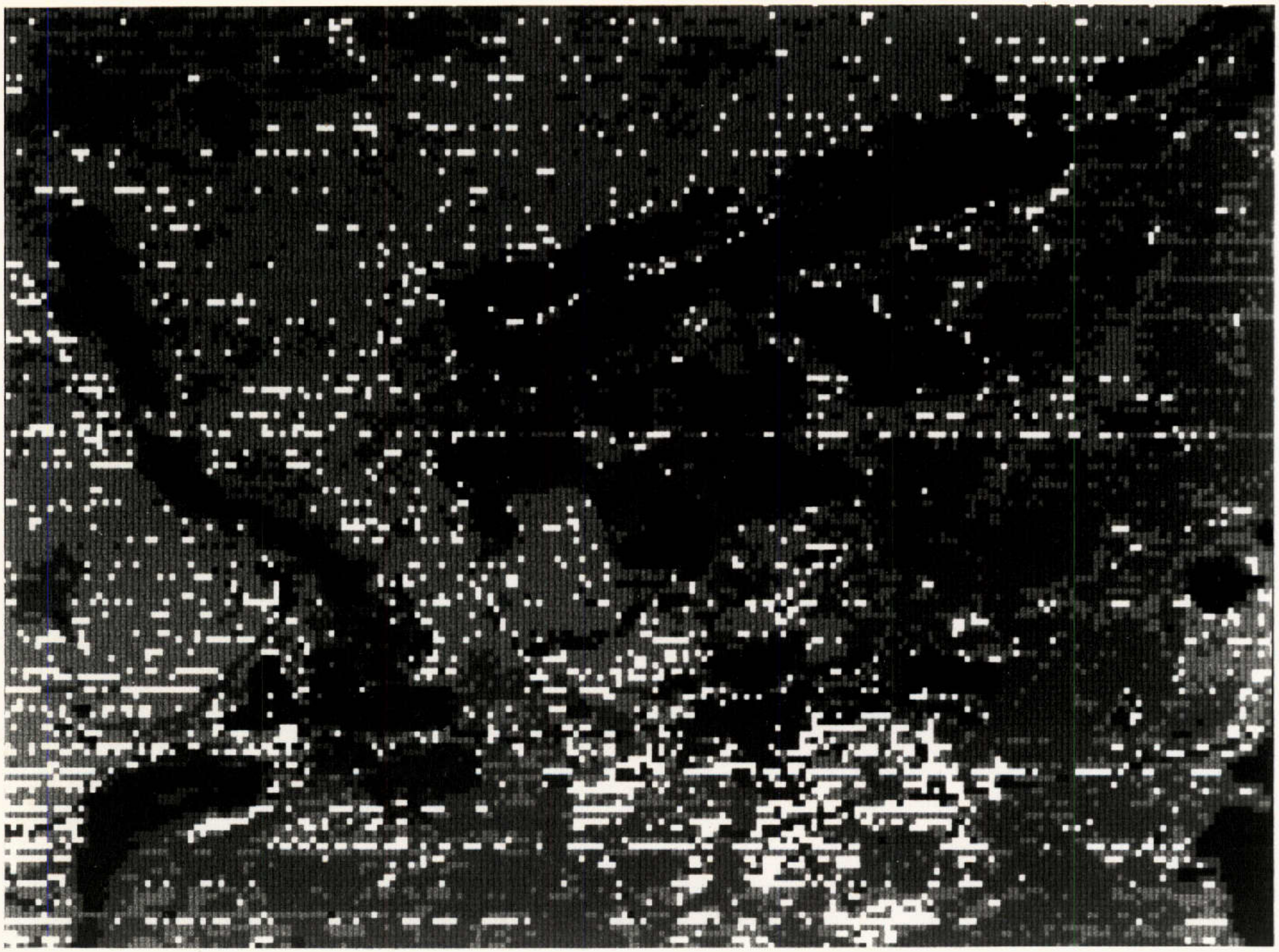


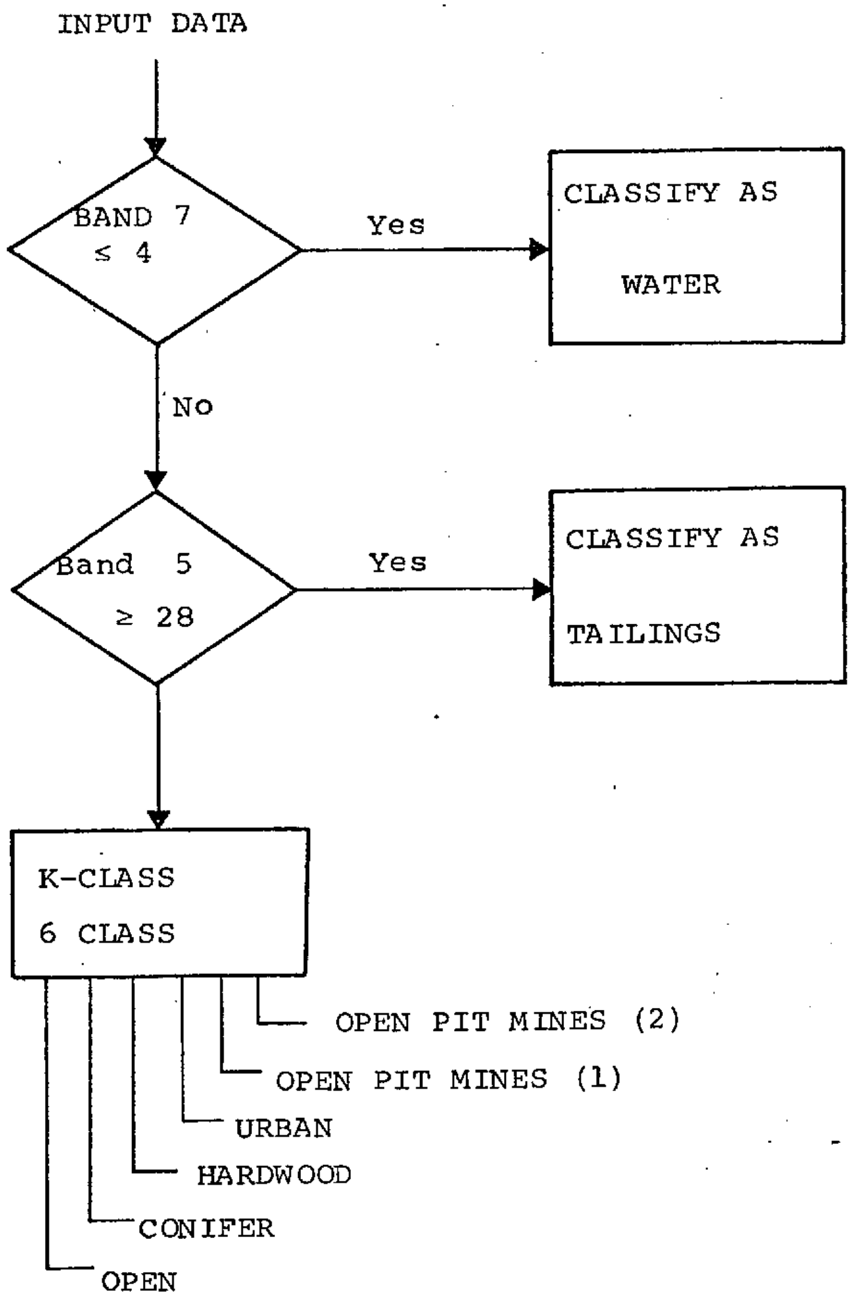
FIGURE 10 PINEVILLE THEMATIC MAP

This thematic map was generated using a sequential test to separate out water and tailings. Water was isolated by those values less than 5 on band 7. Tailings were determined by those values greater than 27 on band 5. The other six classes were classified with K-Class based on all four MSS bands. Open pit mines appeared to have a bi-modal distribution and were broken down into two classes. A block diagram of the classification process is shown in Figure 11.

A comparison can be made between the computer generated thematic map in Figure 10 and the ground truth map shown in Figure 12. The two maps are very similar, indicating that most of the classes are detected accurately. As shown on the thematic map, the class "Open" appears to have the greatest number of errors. The scattered single "Open" indications in the thematic map in Figure 10 may actually be isolated Open areas in the Forested region. One does, however, expect a larger error in this class, as seen in the confusion matrix in Table IV. One reason for the large error in this class may be because of the relatively small number of training samples.

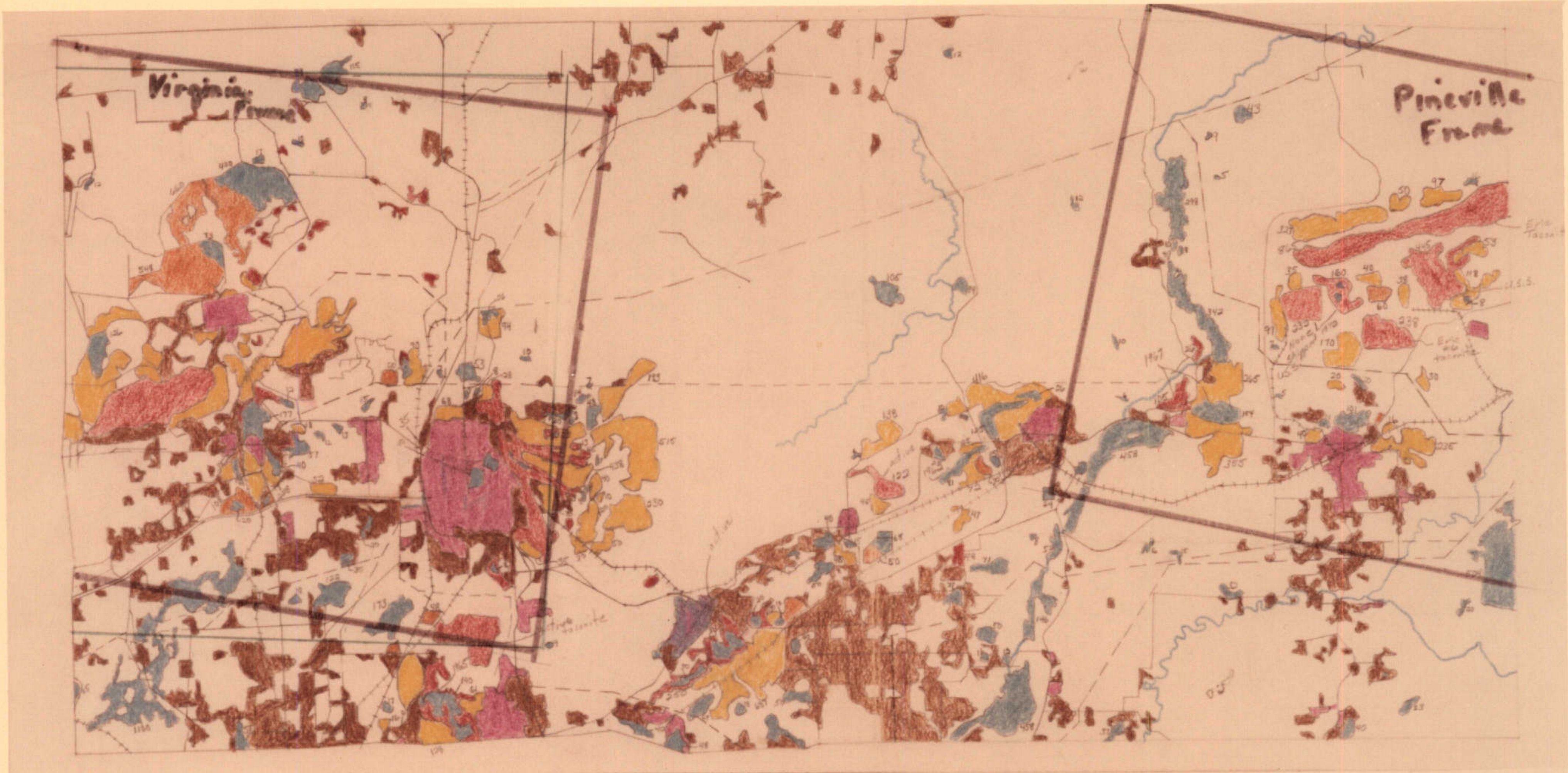
Data from channel 7 of 1129-16320, a November 29, 1972 coverage, was added to the September coverage in an effort to improve classification accuracy. With five bands of multispectral data, a classification accuracy of 89% was obtained for the four channels 1, 3, 4, 6. This compares with 87.5% obtained with the single September coverage only.

The next experiment performed was the application of the weights obtained over Pineville to the Virginia area. The three sets of weights were used for the different channels; that is, a set for scan lines 1, 3, 4, 6, a set for scan line 2 and a set for scan line 5. Figure 13 shows the thematic map generated over Virginia based on the Pineville weights. The confusion matrix for Virginia was derived by selecting random samples from the image and comparing the actual class of that point with the classifier assignment. The confusion matrix obtained in this manner is shown in Table VI.



PINEVILLE 8 CLASS MAPPING

Figure 11. Sequential Classification Procedure



VIRGINIA AND PINEVILLE

GROUND TRUTH

FIGURE 12

BLUE-WATER

BROWN-OPEN, AGRICULTURE

RED-OPEN PIT MINE

YELLOW - TAILINGS PILES

MAGENTA - URBAN

BLANK - FOREST, MARSH

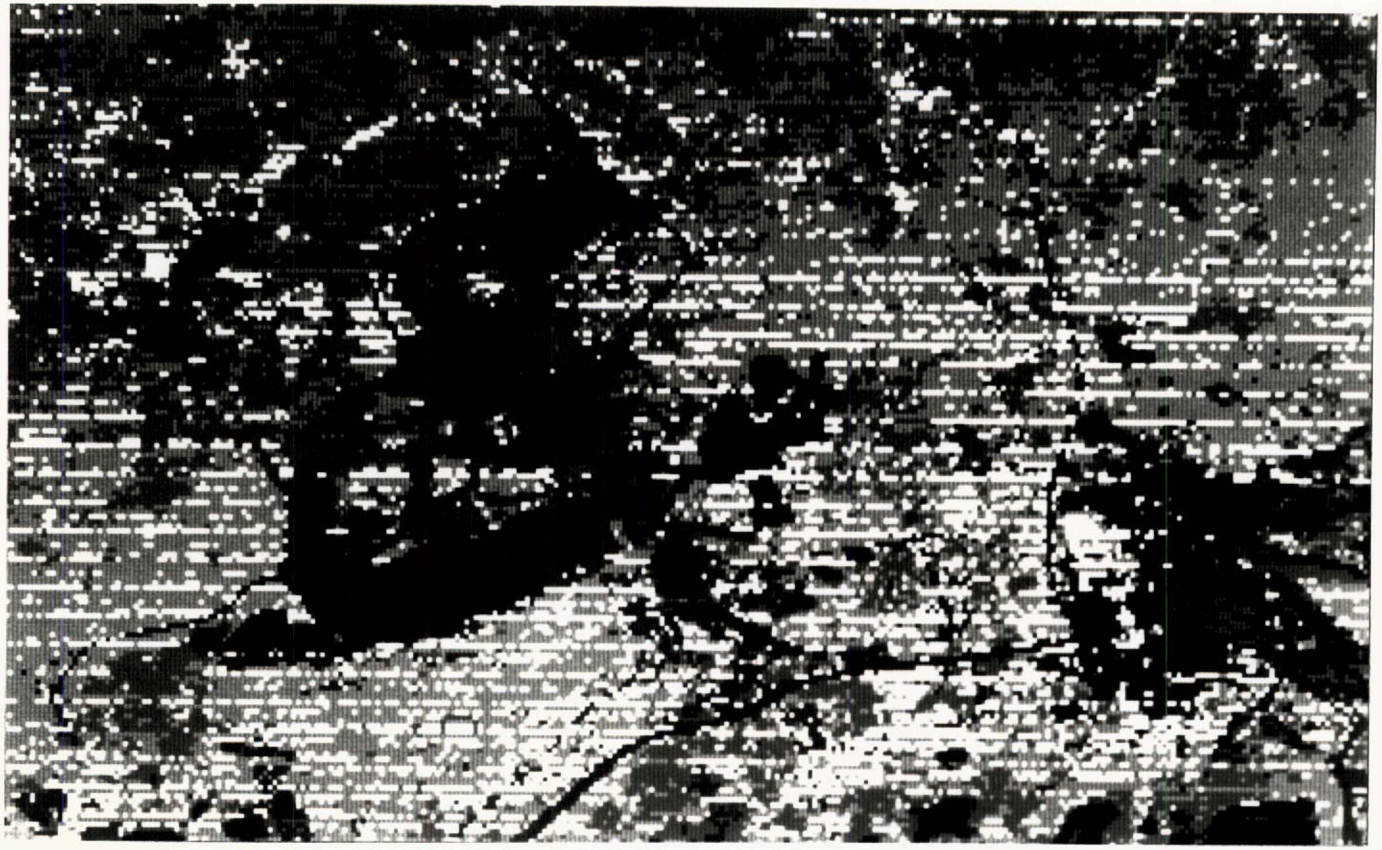


FIGURE 13 VIRGINIA THEMATIC MAP

Table VI.

Virginia Confusion Matrix

Ground truth \ Classifier Assign.	F	R	O	M	W	T
Forest	3745	26	31	6	1	6
Residential	84	155	94	31	1	39
Open	376	30	484	15	0	16
Open Pit Mine	42	9	17	203	2	85
Water	11	0	1	1	138	5
Tailings	21	22	10	57	5	198

The overall classification accuracy of 82.4% is 5.2% less than the overall classification accuracy over Pineville. This indicates that performance does not drop appreciably when the weights obtained from a training set are used for a test set which may be an area 20 miles away. A second evaluation of the accuracy of classification can be made by comparing the computer generated thematic map in Figure 13 with the ground truth map in Figure 12.

Ramsey County Metropolitan Area

The final area considered in this study consisted of Ramsey County, which contains the city of St. Paul, Minnesota. Data for this area was obtained from the ERTS coverage 1075-16321, an October 6th, 1972 cloud-free

overpass. The four multispectral bands were used for features in performing automatic classification. In addition, textural features were used from bands 5 and 7. Severe banding was experienced on Band 4 because channel 2 provided a very low signal output. To accomodate this anamoly, two sets of classification weights were utilized; one set was derived from all four bands but excluded channel 2. The second set used channel 2 from bands 5, 6 and 7 to fill in the missing data.

Three experiments were performed as follows:

1. Using multispectral data only.
2. Using multispectral plus texture from band 7.
3. Using multispectral plus texture from bands 5 and 7.

Four classes were delineated--Water, Residential, Open and Urban. These classes were selected on the basis of availability of training data. The two classes Agriculture and Forest are included in the Open class, since very few samples of this type are available in Ramsey County.

The multispectral features consisted of the four MSS bands (or three for channel 2). Texture was added to these features. Texture was obtained from the Slant Transform coefficients obtained from a 4 x 4 array of data points. When only multispectral data is used, the classification accuracy is about 94%. This performance is obtained using the same data set for training and testing. By adding texture, the performance increased to 100%; however, this again is based on the same training set and test set.

A more realistic evaluation can be obtained by comparing the thematic maps generated by the various procedures. The thematic map obtained using multispectral data alone is shown in Figure 14. The four classes Urban, Water, Open and Residential are plotted from dark to light.

FIGURE 14 RAMSEY COUNTY THEMATIC MAP USING MSS DATA ONLY

When adding texture from band 7, the thematic map shown in Figure 15 was obtained. This map was made by making class assignments on individual pixels; however, the texture portion of the feature was constant over the 4 x 4 array. This accounts for the block like appearance of the thematic map. Most of the block sizes are 4 x 4.

The actual ground truth for Ramsey County is shown in Figure 16. Although this map is 5 years old, the general location of these four classes is accurate. In selecting the training set, updated land use maps were used. These were done by Joe Gibson of the University of Minnesota Department of Geography.

IV. Conclusions

It has been shown that automatic classification of broad classes similar to Anderson Level One can be performed with an accuracy better than 90%. Multi-temporal coverage is very helpful for improving classification accuracy. An improvement of 1.5% was obtained over Pineville by adding a November coverage to a September coverage. Texture is also useful for improving classification accuracy, an improvement of 6% was noted in the metropolitan area study.

Automatic classification using digital data can distinguish more classes and smaller areas than manual photointerpretation on imagery. For example, it has been shown that water bodies of less than an acre size can be detected from the digital area.

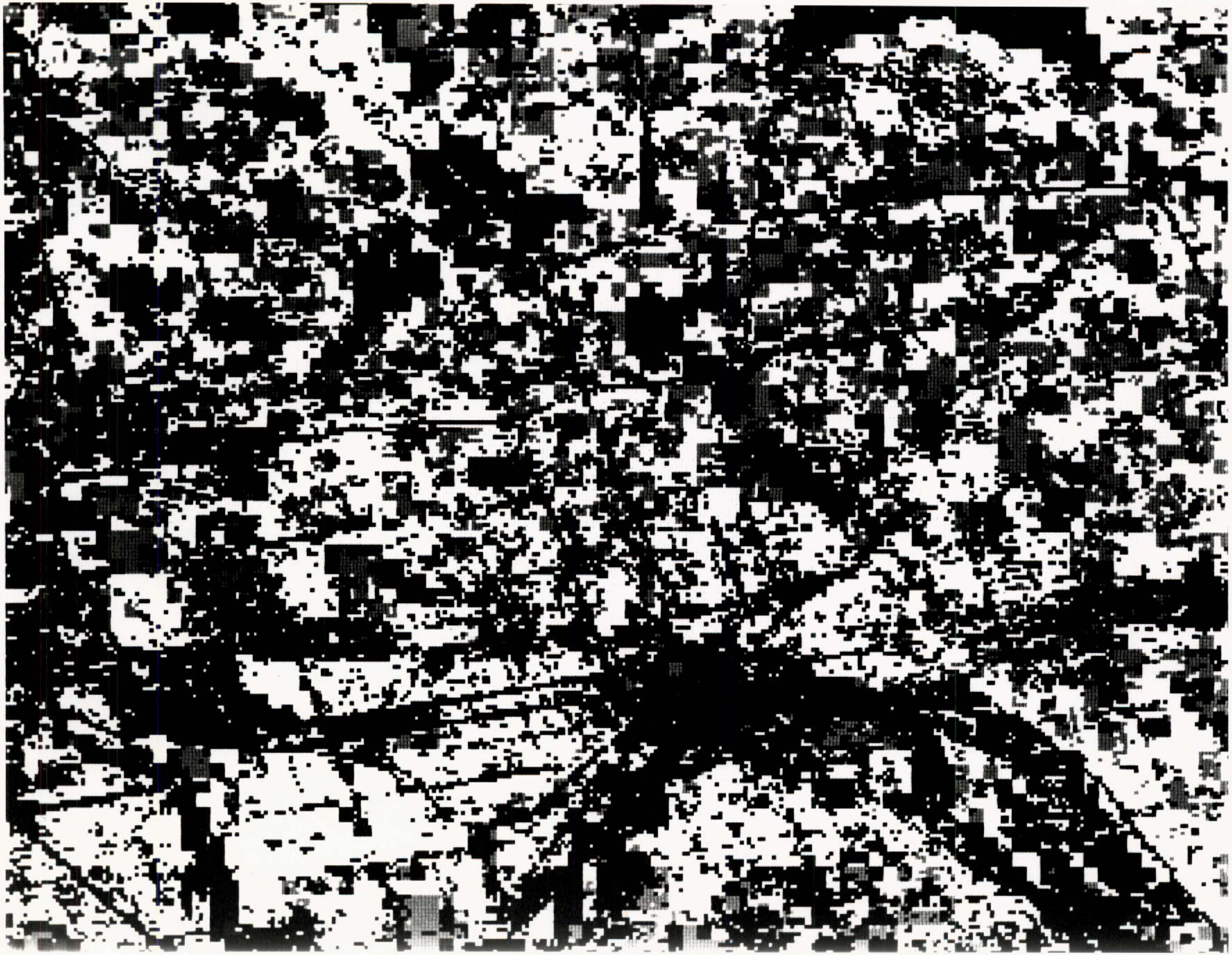


FIGURE 15 RAMSEY COUNTY THEMATIC MAP USING DENSITY AND TEXTURE

FIGURE 16 RAMSEY COUNTY GROUND TRUTH